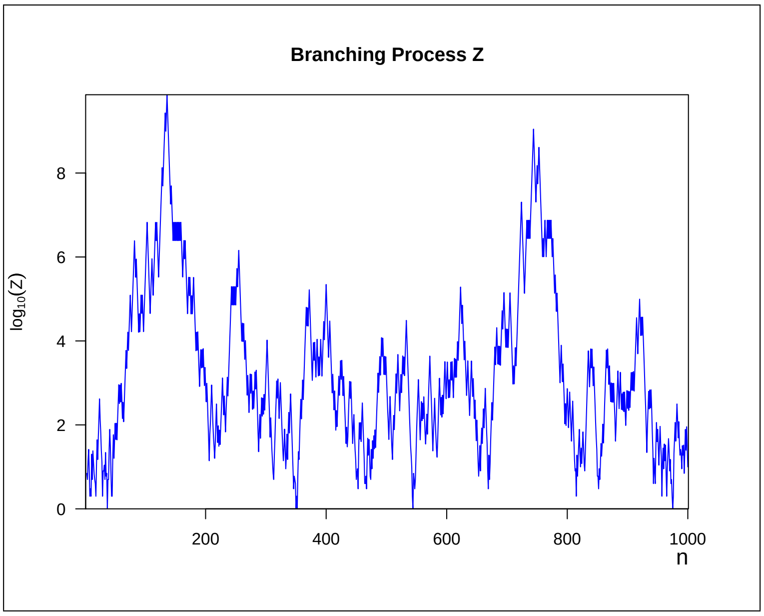
Branching Processes in Random Environment

Dissertation

zur Erlangung des Doktorgrades
der Naturwissenschaften

vorgelegt beim Fachbereich 12, Informatik und Mathematik
der Johann Wolfgang Goethe-Universität
in Frankfurt am Main

von
Christian Böinghoff
aus Erlenbach am Main



Frankfurt am Main 2010
(D30)

vom Fachbereich 12, Informatik und Mathematik der

Johann Wolfgang Goethe-Universität als Dissertation angenommen.

Dekan: Prof. Dr. Tobias Weth

Gutachter: Prof. Dr. Götz Kersting , Prof. Dr. Anton Wakolbinger und Prof. Dr. Nina Gantert

Datum der Disputation: 28.02.2011

Contents

Acknowledgment

The following PhD-thesis emerged from a joint project with scientists from Germany, Russia and France. I would also like to express my gratitude to the German Research Foundation (DFG) and the Russian Foundation of Basic Research for financial support (Grant DFG-RFBR 08-01-91954). It is a pleasure to express gratitude to all people involved in different parts of the project.

First of all, I am indebted to Götz Kersting[1] who supervised this PhD, provided many helpful ideas and always had time to discuss questions with me. His guidance into the world of branching processes has been essential for this thesis.

Secondly, I am very grateful to Vincent Bansaye[2] for the fruitful and enduring collaboration, as well as for invitations to the École Polytechnique and his cordial hospitality.

On the Russian side, I would like to thank Elena Dyakonova, Vladimir Vatutin and Valery Afanasyev[3] for the good cooperation, many fruitful discussions, the invitations to the Steklov Institute in Moscow, and in particular for their extraordinary hospitality during two stays in Moscow.

I am also grateful to Vitali Wachtel[4] for many interesting discussions on BPREs, as well as for quite practical help with language issues during the time in Moscow.

I would also like to thank Nina Gantert[5] for fruitful discussions on the large deviations of BPREs in the quenched approach, as well as for the invitation to Münster.

During my work in Frankfurt, many people helped me with mathematical, technical or organzational questions. Among them, I'd especially like to thank our two secretaries, Nicole Götting and Anna Weiglhofer for their technical support. I'd also like to thank the professors Hermann Dinges, Ralph Neininger, Gaby Schneider and Anton Wakolbinger for many interesting discussions, especially in the vivid stochastic seminary. I'd also like to thank the other PhD students from the stochastic group for the interesting exchange of ideas and the good working atmosphere: Margarete Knape, Max Stroh, Henning Sulzbach and, from statistics, Markus Bingmer who provided very valuable help with a few R problems. I am grateful to Brooks Ferebee for many useful remarks after talks in our seminary and for providing support with subtleties of the English language.

I'd also like to thank my family for their enduring support, especially Andreas for carefully looking for any typos in my PhD-thesis and Ute for carefully reading the pre-final version of the thesis, her advice and many helpful remarks.

[1] Goethe University, Frankfurt/Main
[2] École Polytechnique, Palaiseau
[3] all from Steklov Institute, Moscow
[4] LMU, München
[5] Münster University

Zusammenfassung

In der folgenden Arbeit werden Eigenschaften von ***Verzweigungsprozessen in zufälliger Umgebung*** (engl. *Branching processes in random environment*, kurz ***BPREs***) untersucht. Das Modell geht auf [SW69] und [AK71] zurück. Ein BPRE ist ein einfaches mathematisches Modell für die Entwicklung einer Population von apomiktischen[6] Individuen in diskreter Zeit, wobei die Umgebungsbedingungen einen Einfluß auf den Fortpflanzungserfolg der Individuen haben. Dabei wird angenommen, dass die Umgebungsbedingungen in den einzelnen Generationen zufällig sind, und zwar unabhängig und identisch verteilt von Generation zu Generation. Man denke z.B. an eine Population von Pflanzen mit einem einjährigen Zyklus, die in jedem Jahr anderen Witterungsbedingungen ausgesetzt sind, wobei angenommen wird, dass diese sich unabhängig und identisch verteilt ändern.

Genauer bezeichnen wir eine unendliche Folge von unabhängig, identisch verteilten Zufallsvariablen $Q_1, Q_2, \ldots$, die Werte im Raum Δ aller Wahrscheinlichkeitsverteilungen auf $\mathbb{N}_0$ annehmen, als eine ***Umgebung*** $\Pi = (Q_1, Q_2, \ldots)$. Ein BPRE wird dann wie folgt definiert:

Definition. *Sei $\Pi = (Q_1, Q_2, \ldots)$. Dann bezeichnen wir $(Z_n)_{n\in\mathbb{N}_0}$ als **Verzweigungsprozess in zufälliger Umgebung**, falls für alle $z, k \in \mathbb{N}_0$ die Populationsgröße Z_k in Generation k, gegeben $Z_{k-1} = z$ und gegeben $\Pi = (q_1, q_2, \ldots)$, wie die Summe von z-vielen unabhängig, identisch verteilten Zufallsvariablen verteilt ist, d.h.:*

$$\mathcal{L}(Z_k | Z_{k-1} = z, \Pi = (q_1, q_2, \cdots)) \;=\; \mathcal{L}(\xi_1 + \cdots + \xi_z) \quad ,$$

wobei $\xi_1, \xi_2, \ldots, \xi_z$ unabhängige Zufallsvariablen mit Verteilung q_{k-1} sind.

Als Hilfsmittel definiert man die *zugehörige Irrfahrt*.

Definition. *Seien $\Pi = (Q_1, Q_1, \ldots)$ eine Umgebung und*

$$X_n \;:=\; \log \sum_{y=0}^{\infty} y Q_n(\{y\}), \qquad n \geq 1 \;.$$

*Die Irrfahrt $S = (S_0, S_1, \ldots)$ mit Anfangszustand $S_0 = 0$ und Zuwächsen $X_n = S_n - S_{n-1}$, $n \geq 1$, heißt **zugehörige Irrfahrt** für den Prozess $(Z_n)_{n\in\mathbb{N}_0}$.*

Die zugehörige Irrfahrt bestimmt den Erwartungswert des Prozesses, bedingt auf die Umgebung:

$$\mathbb{E}[Z_n | Z_0, \Pi] \;=\; Z_0 \; e^{S_n} \qquad \text{f.s.}$$

Verzweigungsprozesse in zufälliger Umgebung werden, ähnlich wie gewöhnliche Galton-Watson Prozesse, in superkritische ($\mathbb{E}[X] > 0$), kritische ($\mathbb{E}[X] = 0$) und subkritische Prozesse ($\mathbb{E}[X] < 0$) unterteilt. Kritische und subkritische Prozesse sterben f.s. aus (siehe Kapitel 1).

Bereits in den Arbeiten [Koz76] und [Afa80] wird ein interessantes Verhalten von BPREs im subkritischen Fall beschrieben, zunächst jedoch nur für Nachkommenverteilungen mit gebrochen-linearen Erzeugendenfunktionen. In diesem Fall lässt sich die Erzeugendenfunktion von Z_n, bedingt auf die Umgebung, explizit berechnen. Im subkritischen Fall gibt es drei verschiedene Regime von Verzweigungsprozessen, die sich in der Asymptotik der Überlebenswahrscheinlichkeit und im Verhalten des Prozesses, bedingt auf Überleben (d.h. bedingt auf $\{Z_n > 0\}$, $n \in \mathbb{N}$), unterscheiden. In späteren Arbeiten, z.B. [GKV03], [AGKV05b], [AGKV05a] und [ABKV10] wird dies detailliert beschrieben und unter schwachen Voraussetzungen an die Nachkommenverteilungen und die Verteilung von X gezeigt. Man unterscheidet den schwach subkritischen ($\mathbb{E}[Xe^X] > 0$), den intermediär subkritischen ($\mathbb{E}[Xe^X] = 0$) und den stark subkritischen Fall ($\mathbb{E}[Xe^X] < 0$). Einige bekannte Resultate der letzten Jahre werden in Kapitel 2 vorgestellt.

In [AGKV05b] wird gezeigt, dass Z im stark subkritischen Fall, bedingt auf Überleben, zu allen Zeiten klein bleibt und im Grenzwert als Markovkette beschrieben werden kann. Die Überlebenswahrscheinlichkeit $\mathbb{P}(Z_n > 0)$ fällt exponentiell schnell ab, und zwar mit Ordnung $\mathbb{E}[e^X]^n$.

Der schwach subkritische Fall wird in [ABKV10] beschrieben. Auch in diesem Fall fällt $\mathbb{P}(Z_n > 0)$ exponentiell schnell ab, ist aber von derselben Ordnung wie $\mathbb{P}(\min\{S_0, \ldots, S_n\} \geq 0)$. Unter geeigneten

[6] d.h. sich ungeschlechtlich fortpflanzenden

Voraussetzungen konvergiert der mit dem Erwartungswert skalierte Verzweigungsprozess, bedingt auf Überleben, in Verteilung gegen eine Exkursion einer Brown'schen Bewegung. Dies bedeutet, dass Z_k, bedingt auf $\{Z_n > 0\}$, sowohl für k nahe Null als auch für k nahe n beschränkt bleibt. Dazwischen nimmt Z_k sehr große Werte an und folgt seinem Erwartungswert, bis auf eine zufällige Konstante, auf völlig deterministische Art und Weise. Hier kommt das starke Gesetz der großen Zahl zum Tragen.

Die Untersuchung von intermediär subkritischen Verzweigungsprozessen, bedingt auf Überleben, ist einer der Hauptteile dieser Dissertation. Aufgrund der Bedingung

$$\mathbb{E}[Xe^X] \;=\; 0$$

bietet es sich für die Beweise an, einen Maßwechsel durchzuführen, unter welchem die zugehörige Irrfahrt rekurrent wird. Hierzu definieren wir das Maß $\mathbf{P}$ mit Erwartungswert $\mathbf{E}$ für messbare und beschränkte Funktionen $\Phi : \Delta^n \times \mathbb{R}^n \to \mathbb{R}$ durch

$$\mathbf{E}\left[\Phi(Q_1,\ldots,Q_n,Z_0,\ldots,Z_n)\right] \;=\; \gamma^{-n}\mathbb{E}\left[\Phi(Q_1,\ldots,Q_n,Z_0,\ldots,Z_n)e^{S_n-S_0}\right]$$

mit

$$\gamma \;=\; \mathbb{E}\left[e^X\right] \;.$$

S ist unter $\mathbf{P}$ rekurrent, d.h. $\mathbf{E}[X] = 0$. Wir nehmen an, dass die Verteilung von X folgende Regularitätsbedingung erfüllt:

Annahme 0.1. *Die Verteilung von X hat bzgl. $\mathbf{P}$ endliche Varianz oder, allgemeiner, liegt im Konvergenzbereich einer strikt stabilen Verteilung mit Index $\alpha \in (1, 2]$. Zudem sei sie nicht-gitterartig.*

Außerdem muss eine gewisse Regularität der Nachkommenverteilungen vorausgesetzt werden. Sie betrifft das sogenannte standardisierte zweite Moment von Q. Dazu definiert man

$$\zeta(a) \;=\; \sum_{y=a}^{\infty} y^2 Q(\{y\}) \Big/ m(Q)^2, \quad a \in \mathbb{N} \;.$$

mit $m(Q) = \sum_{y=0}^{\infty} yQ(\{y\})$.

Annahme 0.2. *Es existieren Konstanten $0 < \epsilon < \infty$ und $a \in \mathbb{N}$, so dass*

$$\mathbf{E}[(\log^+ \zeta(a))^{\alpha+\epsilon}] < \infty \;.$$

Wie in der Arbeit später ausführlich erklärt wird, erfüllt eine große Klasse von Verteilungen die obige Bedingung.

Sei

$$\tau_n \;:=\; \min\left\{0 \le k \le n | S_k = \min\{S_0,\ldots,S_n\}\right\} \tag{1}$$

der Zeitpunkt des ersten Minimums von $(S_0, \ldots, S_n)$.

Der folgende Satz wird bereits in [Vat04] gezeigt, jedoch unter etwas stärkeren Voraussetzungen.

Satz 0.1. *Unter den Annahmen 0.1 und 0.2 gilt*

$$\mathbb{P}(Z_n > 0) \;\sim\; \gamma^n \theta \mathbf{P}(\tau_n = n)$$

für ein $0 < \theta < \infty$.

Der nächste Satz beschreibt die Anzahl der Zeiten – bedingt aufs Überleben des Prozesses – zu denen nur noch genau ein Individuum lebt. Dazu nehmen wir an, dass es mit positiver Wahrscheinlichkeit Verteilungen gibt, unter welchen ein Individuum mit positiver Wahrscheinlichkeit keine oder genau einen Nachkommen haben kann:

Annahme 0.3.

$$\mathbb{E}\big[Q(\{1\})Q(\{0\})\big] \;>\; 0 \;.$$

Die Anzahl der Zeitpunkte, zu denen nur noch genau ein Individuum lebt ist dann von der gleichen Ordnung wie die Anzahl der strikt absteigenden Leiterpunkte einer rekurrenten Irrfahrt, bedingt auf $\{\tau_n = n\}$.

Satz 0.2. *Unter den Annahmen 0.1 bis 0.3 gibt es eine schwach variierende Folge* $b_1, b_2, \ldots$, *so dass*

$$\mathbb{E}\Big[\sharp\{k|Z_k = 1\}\Big|Z_n > 0\Big] \quad = \quad \Theta(b_n\ n^{1-\rho})$$

gilt.

Dabei bezeichnet $x_n = \Theta(y_n)$, dass die Folge x_n für $n \to \infty$ von der gleichen Ordnung wie y_n ist. Genauer gesagt gibt es Konstanten $c_1, c_2 \in \mathbb{R}^+$, so dass

$$c_1 \ \leq\ \liminf_{n\to\infty} \frac{x_n}{y_n} \quad \leq \quad \limsup_{n\to\infty} \frac{x_n}{y_n} \ \leq\ c_2\ .$$

Wie bereits durch Satz 0.1 angedeutet wird, konvergiert die zugehörige Irrfahrt, bedingt aufs Überleben des Prozesses, gegen einen Lévy-Prozess, der darauf bedingt ist, sein Minimum am Ende anzunehmen.

Satz 0.3. *Unter Annahmen 0.1 und 0.2 gilt für* $n \to \infty$,

$$\mathcal{L}\Big(\big(S_{\lfloor nt\rfloor}/a_{\lfloor nt\rfloor}\big)_{0\leq t\leq 1}\big|Z_n > 0\Big) \quad \xrightarrow{d} \quad \mathcal{L}(L^-)\ ,$$

in Verteilung bzgl. der Skorohod Metrik, wobei L^- *einen Lévy-Prozess bezeichnet, der darauf bedingt ist, sein Minimum am Ende anzunehmen.*

Im intermediär subkritischen Fall zeichnet sich also folgendes Bild ab: Der Prozess überlebt typischerweise in Umgebungen in denen sich die zugehörige Irrfahrt wie eine rekurrente Irrfahrt verhält, die darauf bedingt ist, ihr Minimum am Ende anzunehmen. Das heißt, dass Überleben nicht allein dadurch realisiert wird, dass der Prozess einer außergewöhnlich 'günstigen' Umgebung[7] ausgesetzt ist (wie im schwach subkritischen Fall) sondern durch ungewöhnlich hohe Nachkommenzahlen innerhalb einer 'ungünstigen' Umgebung[8]. Eine rekurrente Irrfahrt, bedingt darauf ihr Minimum am Ende anzunehmen, kann lange Exkursionen zwischen den absteigenden Leiterpunkten besitzen. In diesen Perioden folgt der Prozess seinem Erwartungswert.

Im zweiten Teil der Arbeit werden große Abweichungen behandelt. Die Wahrscheinlichkeit, dass der Prozess außergewöhlich große Werte annimmt, fällt exponentiell schnell ab. In dieser Arbeit wird die Ratenfunktion $\psi : \mathbb{R} \to \mathbb{R}^+ \cup \{\infty\}$ bestimmt, die

$$\mathbb{P}(Z_n > e^{\theta n}) \ = \ e^{-\psi(\theta)n+o(n)}$$

erfüllt. Für unsere Untersuchung nehmen wir an, dass eine nichtentartete Ratenfunktion für die zugehörige Irrfahrt existiert, was durch die sogenannte *rechtsseitige Cramér-Bedingung* sichergestellt wird. Es wird gefordert, dass die zugehörige Irrfahrt endliche exponentielle Momente besitzt:

Annahme 0.4. *Es existiert ein* $s > 0$, *so dass die momentenerzeugende Funktion endlich ist:*

$$\varphi(s) \quad := \quad \mathbb{E}\left[e^{sX}\right] \ < \ \infty\ .$$

Insbesondere existiert $\mathbb{E}[X] \geq -\infty$.

Für subkritische BPREs ist bereits das Überleben des Prozesses ein Ereignis mit exponentiell schnell abfallender Wahrscheinlichkeit. Ein einfaches Subadditivitätsargument liefert allgemein die Existenz des Grenzwertes

$$\gamma \quad := \quad \lim_{n\to\infty} -\frac{1}{n} \log \mathbb{P}(Z_n > 0)\ ,$$

[7]'günstige' Umgebung bedeutet $\min\{S_0, \cdots, S_n\} \geq 0$

[8]'ungünstige' Umgebung bedeutet, dass das $\min\{S_0, \cdots, S_n\}$ klein ist

wobei $0 \leq \gamma < \infty$.

Sei Λ die Ratenfunktion der zugehörigen Irrfahrt, d.h.

$$\mathbb{P}(S_n \geq \theta n) \ = \ e^{-\Lambda(\theta)n+o(n)}$$

mit

$$\Lambda(\theta) \quad := \quad \sup_{s\geq 0} \left\{ s\theta - \log\varphi(s) \right\} \ .$$

Wie sich herausstellt, hängt für eine große Klasse von Nachkommenverteilungen (diejenigen, bei denen alle Momente endlich sind) die Ratenfunktion ψ nur von γ und Λ ab.

Von Bedeutung ist die Funktion $\Gamma : \mathbb{R}^+ \to \mathbb{R} \cup \{\infty\}$. Für alle $\theta \geq 0$ ist Γ durch

$$\Gamma \ := \ \inf_{0<t\leq 1} \left\{ t\gamma + (1-t)\Lambda(\theta/(1-t)) \right\} \tag{2}$$

definiert.

In Kapitel 4.2 wird zunächst der Fall von Nachkommenverteilungen behandelt, deren Tails geometrisch schnell abfallen, was durch folgende Annahme sichergestellt wird.

Annahme 0.5. *Es existieren Konstanten $k_0 \in \mathbb{N}_0$, $0 \leq a < b$ und $c > 0$, so dass Q f.s. Werte in der Menge aller Wahrscheinlichkeitsverteilungen $\mathcal{A} \subset \Delta$ mit der folgenden Eigenschaft annimmt:*
Falls R Verteilung $\mathcal{P}$ und Erwartungswert $\mathcal{E}[R] = m$ hat, so gilt

$$\mathcal{E}\left[(R-j)^+\right] \quad \leq \quad c\ m\ \left(\frac{a+m}{b+m}\right)^{j-k_0} \ , \quad j \geq k_0. \tag{3}$$

Unter dieser Voraussetzung gilt $\psi = \Gamma$, d.h.

Satz 0.4. *Unter Annahmen 0.4 und 0.5 gilt für jedes $\theta \geq 0$*

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}(Z_n \geq e^{\theta n}) \quad \leq \quad -\Gamma(\theta) \quad ,$$
$$\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}(Z_n > e^{\theta n}) \quad \geq \quad -\Gamma(\theta+) \ .$$

In Kapitel 4.3 werden Nachkommenverteilungen mit schweren Tails behandelt.

Die folgende Annahme stellt sicher, dass die Tails der Nachkommenverteilungen, gleichmäßig über alle Umgebungen, mindestens mit Exponent $\beta \in (1, \infty)$ abfallen:

Annahme 0.6. *Es existiert eine Konstante $0 < d < \infty$, so dass Q f.s. Werte im Raum der Wahrscheinlichkeitsverteilungen $\mathcal{A} \subset \Delta$ mit der folgenden Eigenschaft annimmt:*
Falls R Verteilung $\mathcal{P}$ und Erwartungswert $\mathcal{E}[R] = m$ hat, so gilt für alle $z > 0$

$$\mathcal{P}(R > z \ | R > 0) \quad \leq \quad d\ (m \wedge 1)\ z^{-\beta} \quad \textit{f.s.}$$

Die Ratenfunktion ψ hängt dann nicht nur von γ und Λ ab, sondern auch von β. Sie ist durch

$$\psi(\theta) \ = \ \psi_{\gamma,\beta,\Lambda}(\theta) \ := \ \inf_{t\in[0,1],s\in[0,\theta]} \Big\{ t\gamma + \beta s + (1-t)\Lambda((\theta-s)/(1-t)) \Big\} \tag{4}$$

definiert.

In Kapitel 4.3 wird dann folgender Satz gezeigt:

Satz 0.5. *Falls ein $\beta \in (1, \infty)$ existiert mit $\log(\mathbb{P}(Z_1 > z))/\log(z) \stackrel{z\to\infty}{\longrightarrow} -\beta$ und zusätzlich Annahme 0.6 für dieses β erfüllt ist, so gilt für jedes $\theta \geq 0$*

$$\frac{1}{n} \log(\mathbb{P}(Z_n \geq e^{\theta n})) \stackrel{n\to\infty}{\longrightarrow} -\psi(\theta).$$

Bemerkung. *Die erste Annahme in Satz 0.5 stellt sicher, dass mit positiver Wahrscheinlichkeit Nachommenverteilungen mit schweren Tails auftreten, deren Tails mit Exponent β abfallen.*

Mithilfe des obigen Satzes lässt sich Satz 0.4 leicht verallgemeinern:

Satz 0.6. *Falls Annahme 0.6 für jedes $\beta > 0$ erfüllt ist, so gilt für jedes $\theta \geq 0$*

$$\begin{aligned} \limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}(Z_n \geq e^{\theta n}) &\leq -\Gamma(\theta) \quad , \\ \liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}(Z_n > e^{\theta n}) &\geq -\Gamma(\theta+) \ , \end{aligned}$$

wobei $\Gamma(\theta) = \inf_{t\in[0,1]}\{t\gamma + (1-t)\Lambda(\theta/(1-t))\}$.

In Kapitel 4.3 werden Ereignisse untersucht, deren Wahrscheinlichkeiten exponentiell schnell abfallen und zwar mit einer Rate, die durch die Ratenfunktion ψ bestimmt wird. Diese Ereignisse können auf verschiedenen Wegen realisiert werden. Der exponentielle Abfall wird jedoch durch den 'günstigsten' Weg bestimmt. Im Fall von Verteilungen mit schweren Tails und für große θ besteht diese 'optimale Strategie' daraus, dass zunächst ein Individuum exponentiell viele Nachkommen hat und der Prozess anschließend in einer günstigen Umgebung (d.h. hier, dass die zugehörige Irrfahrt linear wächst) gemäß seinem Erwartungswert wächst.

Im stark subkritischen Fall und für kleine θ besteht die optimale Strategie daraus, zunächst in einer ungünstigen Umgebung (d.h. hier, dass die zugehörige Irrfahrt linear fällt) bis zur Zeit $\lfloor tn \rfloor$, $t \in (0,1)$ nur zu überleben. Der Prozess überlebt zwar, bleibt aber beschränkt. Erst ab dem Zeitpunkt $\lfloor tn \rfloor$ wird eine günstige Umgebung realisiert und der Prozess wächst entsprechend seinem Erwartungswert. Dieser Effekt wird, ebenso wie Sonderfälle, detailliert in Kapitel 4.3.1 beschrieben.

Die Beweise von Satz 0.4 und 0.5 beruhen darauf, als untere Schranke die Wahrscheinlichkeiten entlang einer 'optimalen Strategie' zu maximieren. Für die obere Schranke wird eine Abschätzung der Tailwahrscheinlichkeiten von Z_n, bedingt auf die Umgebung, benötigt. Diese Abschätzung erhält man über die Berechnung und Abschätzung von Ableitungen von Erzeugendenfunktionen.

In Kapitel 4.4 wird die exponetielle Abfallrate der Wahrscheinlichkeit $\mathbb{P}(0 \leq Z_n \leq e^{\theta n})$ für superkritische BPREs ($\mathbb{E}[X] > 0$) im Fall von Nachkommenverteilungen mit gebrochen-linearer Erzeugendenfunktion bestimmt. In diesem Fall kann die Verteilung von Z_n, bedingt auf die Umgebung, explizit berechnet werden.

Annahme 0.7. *Q nimmt $\mathbb{P}-$ f.s. Werte in der Menge aller Wahrscheinlichkeitsverteilungen $\mathcal{A} \subset \Delta$ mit der folgenden Eigenschaft an:*
Sei R eine Zufallsvariable mit Verteilung $\mathcal{P}$. Dann ist für $s \in [0,1]$ die Erzeugendenfunktion durch

$$f_R(s) := \mathcal{E}[s^R] \;=\; 1 - \frac{1-s}{m_R^{-1} + 1/2\ b_R m_R^{-2}(1-s)}$$

gegeben, mit $m_R = \sum_{k=0}^{\infty} k\mathcal{P}(R=k)$ und $b_R = \sum_{k=0}^{\infty} k(k-1)\mathcal{P}(R=k)$. Zusätzlich existieren Konstanten $0 < c_1 < c_2 < \infty$ (gleichmäßig für alle $\mathcal{P}$) so dass

$$c_1 \;<\; 1/2\ b_R m_R^{-2} \;<\; c_2 \ .$$

Zur Vereinfachung wird angenommen, dass die momentenerzeugende Funktion der Zuwächse der Irrfahrt für alle $s \in \mathbb{R}$ endlich ist,

$$\varphi(s) \;:=\; \mathbb{E}[e^{sX}] \;<\; \infty \quad \text{für alle } s \in \mathbb{R}$$

und ungünstige Umgebungen auftreten können, d.h. $\mathbb{P}(X < 0) > 0$.

Definiert man die exponentielle Abfallrate der Wahrscheinlichkeit, dass der Prozess überlebt, aber beschränkt bleibt,

$$\varrho \;:=\; -\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}(Z_n = 1) \ , \tag{5}$$

so ist die gesuchte Ratenfunktionn χ durch

$$\chi(\theta) \quad := \quad \inf_{t\in[0,1)} \left\{ t\varrho + (1-t)\Lambda(\theta/(1-t)) \right\} \tag{6}$$

gegeben. Dann bestimmt χ im Fall geometrischer Nachkommenverteilungen die exponentielle Abfallrate von $\mathbb{P}(0 \leq Z_n \leq e^{\theta n})$:

Satz 0.7. *Es sei* $\mathbb{E}[X] > 0$ *und Annahme 0.7 erfüllt. Dann gilt für alle* $0 < \theta < \mathbb{E}[X]$

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}(1 \leq Z_n \leq e^{\theta n}) \quad = \quad -\chi(\theta)$$

und für alle $\theta \geq \mathbb{E}[X]$,

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}(Z_n \geq e^{\theta n}) \quad = \quad -\chi(\theta) \ (= \Lambda(\theta)) \ .$$

Für kleine θ besteht nun die optimale Strategie, zu Überleben und kleiner als $e^{\theta n}$ zu bleiben darin, bis zur Zeit $\lfloor tn \rfloor$ innerhalb einer günstigen Umgebung beschränkt zu bleiben. Erst anschließend wächst der Prozess entsprechend seinem Erwartungswert.

In Kapitel 4.5 werden große Abweichungen, bedingt auf die Umgebung (engl. *quenched*), untersucht. Außergewöhnlich große Werte können hier nicht mehr über die Umgebung realisiert werden. Für Nachkommenverteilungen mit schweren Tails gilt:

Satz 0.8. *Falls* $\limsup_{z\to\infty} \log \mathbb{P}(Z_1 > z|\Pi, Z_0 = 1)/\log z = -\beta$ *f.s. für ein* $\beta \in (1,\infty)$ *und Annahme 0.6 für dieses* β *erfüllt ist, so gilt für jedes* $\theta \geq (\mathbf{E}[X] \vee 0)$,

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}(Z_n > e^{\theta n}|\Pi) \quad = \quad \begin{cases} -\beta(\theta - \mathbb{E}[X]) & , \ \textit{falls} \ \ \mathbb{E}[X] > 0 \\ -(\beta\theta - \mathbb{E}[X]) & , \ \textit{falls} \ \ \mathbb{E}[X] \leq 0 \end{cases} \qquad \textit{f.s.}$$

Falls alle Nachkommenverteilungen geometrisch beschränkte Tailwahrscheinlichkeiten besitzen, so gilt:

Satz 0.9. *Falls* $\mathbb{P}(Z_1 > e^{\theta}|\Pi) > 0$ *f.s. und Annahme 0.5 erfüllt ist, so gilt für jedes* $\theta \geq (\mathbf{E}[X] \vee 0)$

$$\lim_{n\to\infty} \frac{1}{n} \log(-\log \mathbb{P}(Z_n > e^{\theta n}|\Pi)) \quad = \quad \begin{cases} \theta - \mathbb{E}[X] & , \ \textit{falls} \ \ \mathbb{E}[X] > 0 \\ \theta & , \ \textit{falls} \ \ \mathbb{E}[X] \leq 0 \end{cases} \qquad \textit{f.s.}$$

Die Wahrscheinlichkeit, außergewöhnlich große Werte zu realisieren, ist von kleinerer als exponentieller Ordnung.

Zum Abschluss der Dissertation werden Verzeigungsprozesse in zufälliger Umgebung, bedingt auf Überleben, simuliert. Dazu wird eine Konstruktion nach [Gei99] angewendet. Diese erlaubt es, Galton-Watson Bäume in variierender Umgebung, bedingt auf Überleben, entlang einer Ahnenlinie zu konstruieren. Der Fall von Nachkommenverteilungen mit gebrochen-linearen Erzeugendenfunktionen, auf den wir uns in Kapitel 5 beschränken, erlaubt die explizite Berechnung der benötigten Verteilungen. Als Anwendung von Satz 0.3 können nun intermediär subkritische Verzeweigungsprozesse, bedingt auf Überleben, wie folgt simuliert werden: Zunächst wird die Umgebung zufällig bestimmt, und zwar als Irrfahrt, bedingt darauf ihr Minimum am Ende anzunehmen. Anschließend wird, der Geiger-Konstruktion folgend, ein Verzweigungsprozess in dieser Umgebung, bedingt auf Überleben, simuliert.

Zum Abschluss wird in einem Ausblick auf aktuelle Forschung verwiesen. Im Anhang befinden sich einige technische Resultate.

Chapter 1

Introduction

1.1 Historical remarks

Think of a population of apomictic[9] plants having a one year life cycle. Each year, the weather conditions (the environment) vary, which influences the reproductive success of the plants. Given the climate, all plants reproduce according to the same mechanism. In the simplified model here, the environment is assumed to be independently and identically distributed (i.i.d.). Thus, in each generation, an offspring distribution is chosen at random, independently from one generation to the other. This is the toy model for a branching process in random environment (BPRE) (see Figure 1.1 for an example).

BPREs have first been introduced in [SW69] and [AK71]. Initially, they have mainly been studied under the assumption of i.i.d. offspring distributions which are geometric or, more generally, have generating functions which are linear fractional (see [Koz76], [Afa80]). In recent years, the case of general offspring distributions has attracted attention (compare [GK00], [BGK05], [AGKV05a], [AGKV05b], [VK08], [Ban09], [Afa10], [ABKV10]), as well as special topics like large deviations, e.g. [Koz06], [BB09] and [HL10].

A list of older results is [Tan77], [Tan78], [CT84], [Gui85], [Tan88], [GZ91], [Ham92], [Afa93], [Koz95], [Liu96], [BV97], [DH97a], [Afa97], [VD97],[Afa98], [Afa01a], [VD02], [DGV04], [VD04].

In the next chapter, the mathematical model of a BPRE is described more in detail.

1.2 The model

In this section, the formal definition of the BPRE described in the preceding paragraph is presented and some basic properties of the model are described. By Δ we denote the space of all probability measures on $\mathbb{N}_0 := \{0, 1, 2, 3, \ldots\}$. Equipped with the metric of total variation[10], Δ is a Polish space. Let Q be a random variable taking values in Δ. An infinite sequence $\Pi = (Q_1, Q_2, \ldots)$ of i.i.d. copies of Q is called a ***random environment*** and Q_n the offspring distribution in generation $n - 1$.

Definition. *A process $Z = (Z_0, Z_1, \ldots)$ with values in $\mathbb{N}_0$ is called a **branching process in random environment** Π, if Z_0 is independent of Π and if, given Π, Z is a Markov chain and for every $n \geq 1$, $z \in \mathbb{N}_0$, $q_1, q_2, \ldots \in \Delta$*

$$\mathcal{L}\big(Z_n | Z_{n-1} = z, \Pi = (q_1, q_2, \ldots)\big) \quad = \quad \mathcal{L}(\xi_1 + \cdots + \xi_z) \ , \tag{1.1}$$

where $\xi_1, \xi_2, \ldots$ are i.i.d. random variables with distribution q_n.

Z_n is called the n^{th} generation size.

Fine properties of Z are mainly determined by an auxiliary process, called ***associated random walk***, which depends on the mean offspring number in each generation.

[9] *apomoxis*: Reproduction without fertilization

[10] For $q_1, q_2 \in \Delta$, the total variation metric is defined by $d_{TV}(q_1, q_2) := \sum_{k=0}^{\infty} |q_1(\{k\}) - q_2(\{k\})|$

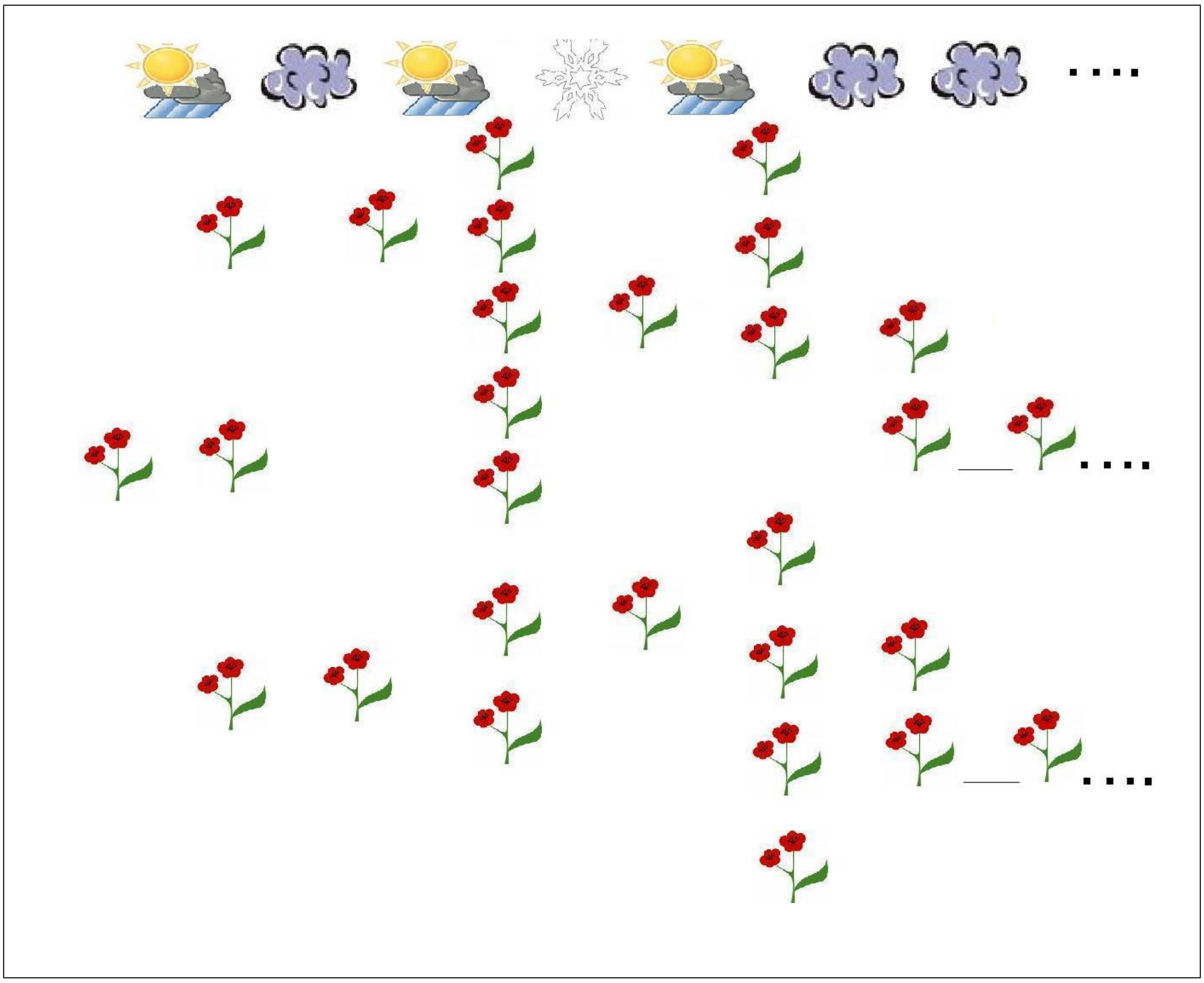


Figure 1.1: Illustration of a BPRE with three possible environmental states (taken from the presentation of Vincent Bansaye, *'Large deviations for branching processes in random environment'*, 10. july 2009, École d'été de Probabilités, Saint-Flour).

Definition. *Set*

$$X_n \quad := \quad \log \sum_{y=0}^{\infty} y \; Q_n(\{y\}), \quad n \geq 1 \; .$$

The random walk $S = (S_0, S_1, \ldots)$ with initial state $S_0 = 0$ and increments $X_n = S_n - S_{n-1}$, $n \geq 1$ is called ***associated random walk*** *for the process $(Z_n)_{n \in \mathbb{N}_0}$.*

Notice that the X_n are i.i.d. copies of the logarithmic mean offspring number

$$X \quad = \quad \log \sum_{y=0}^{\infty} y \; Q(\{y\}) \; ,$$

which is assumed finite a.s. Thus, the conditioned means of Z_n may be written as

$$\begin{aligned} \mathbb{E}[Z_n | Z_0 = z, \Pi] \quad &= \quad z \; e^{X_1} \cdots e^{X_n} \\ &= \quad z \; e^{S_n} \quad \text{a.s.} \end{aligned} \tag{1.2}$$

Compare Figure 1.2 for an example.

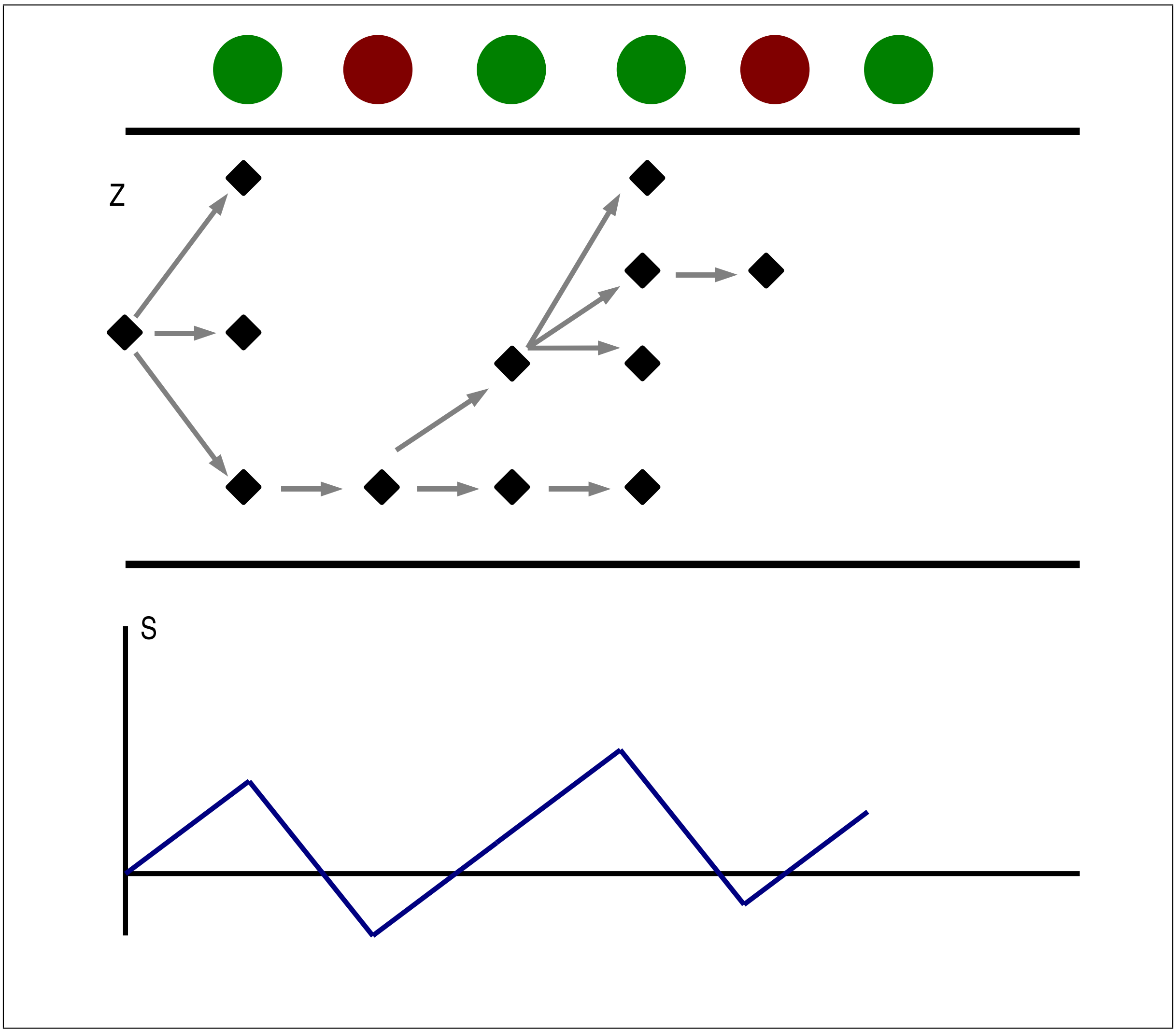


Figure 1.2: Illustration of a BPRE with two possible environmental states (good and bad).

In the theory of classical Galton-Watson processes, three cases are distinguished (see [AK71], p. 8) according to the mean offspring number

$$\begin{array}{lcll} \mathbb{E}[Z_1|Z_0=1] & > & 1 & \text{supercritical case} \\ \mathbb{E}[Z_1|Z_0=1] & = & 1 & \text{critical case} \\ \mathbb{E}[Z_1|Z_0=1] & < & 1 & \text{subcritical case .} \end{array}$$

For BPREs, a different classification is needed. The supercritical, critical and subcritical cases are distinguished according to the drift of the associated random walk (see e.g. [BGK05]). First, if S has positive drift (i.e. $\lim_{n\to\infty} S_n = \infty$ a.s., see [Fel87]), $\mathbb{E}[Z_n|\Pi] \to \infty$ a.s. as n tends to infinity. This is called the ***supercritical*** case. Second, if S has negative drift (i.e. $\lim_{n\to\infty} S_n = -\infty$ a.s.) the process is called ***subcritical***. Finally, if S is an oscillating random walk (meaning $\limsup_{n\to\infty} S_n = \infty$ a.s. and $\liminf_{n\to\infty} S_n = -\infty$ a.s.), the BPRE is called ***critical***.

In the classical works on BPREs ([SW69], [AK71]), it has been assumed that S has finite mean. Then Z is called supercritical, subcritical or critical according as $\mathbb{E}[X] > 0$, $\mathbb{E}[X] < 0$ or $\mathbb{E}[X] = 0$. Recently, the assumption of the existence of $\mathbb{E}[X]$ has been dropped (see [AGKV05b] for the strongly subcritical and [AGKV05a] for the critical case).
A nice survey of the situation in the critical and subcritical cases and an explanation of the heuristics can be found in [BGK05] which we will recall in the sequel.

The classification of BPREs follows the same lines as the distinction of classical Galton-Watson processes, although there are differences which we will explain later. In the critical and subcritical cases, the population becomes extinct with probability 1. This is an immediate consequence of a first moment estimate. For all $m \le n$,

$$\mathbb{P}(Z_n > 0|\Pi) \quad \le \quad \mathbb{P}(Z_m > 0|\Pi) \;\le\; e^{S_m} \qquad \text{a.s.}$$

and thus

$$\mathbb{P}(Z_n > 0|\Pi) \quad \le \quad \min_{m\le n} e^{S_m} \;=\; \exp\big(\min_{m\le n} S_m\big) \qquad \text{a.s.} \tag{1.3}$$

For critical and subcritical BPREs, this implies $\mathbb{P}(Z_n > 0|\Pi) \to 0$ a.s. and thus $\mathbb{P}(Z_n \to 0) = 1$. In contrast to classical Galton-Watson processes, the converse is not always true. Even in the supercritical case, it may happen that the process dies out a.s. (within only few generations) due to random fluctuations, a fact which will be described more in detail in the next section.
Let us shortly explain the heuristics behind the classification of BPREs. For simplicity, assume that $\mathbb{E}[X^2] < \infty$ and that (1.3) gives the correct order of decay of the survival probability (up to a constant), that is

$$\mathbb{P}(Z_n > 0) \;\sim\; c\,\mathbb{E}\big[\exp\big(\min_{m\le n} S_m\big)\big].$$

It then remains to analyze the asymptotic of $\mathbb{E}\big[\exp\big(\min_{m\le n} S_m\big)\big]$. In the critical case, that is for $\mathbb{E}[X] = 0$, there will only be a considerable contribution to the expectation if $\min_{m\le n} S_m$ is close to zero. In the finite variance case, it is well-known that the probability of $\{\min_{m\le n} S_m \ge 0\}$ is of the order $n^{-1/2}$ (see e.g. [GKV03]) and we expect

$$\mathbb{P}(Z_n > 0) \;\sim\; c\,\mathbb{P}\big(\min_{1\le m\le n} S_m \ge 0\big) \;\sim\; c'\,n^{-1/2} \qquad .$$

for some constant $c' > 0$. As it will be detailed in the next section, this result is also true in a more general context and does not require finite variance of the associated random walk.
Introduce

$$\begin{aligned} M_n &:= \max_{1\le j\le n} S_j \\ L_n &:= \min_{1\le j\le n} S_j \;. \end{aligned}$$

By τ_n, we denote the first time, when the minimum of $S_0, \dots, S_n$ is attained

$$\tau_n \quad := \quad \min\{0 \le k \le n | S_k = \min\{S_0, \dots, S_n\}\} \quad n \ge 0 \;. \tag{1.4}$$

Now consider a subcritical BPRE where the associated random walk has negative drift. Then the probability of $\{L_n \ge 0\}$ is exponentially small and it is more complicated to estimate the asymptotic of $\mathbb{E}\big[\exp\big(L_n\big)\big]$. For this, we will use a change of measure and define

$$\beta \quad := \quad \sup\big\{0 \le s < \infty \big| \mathbb{E}\big[Xe^{sX}\big] = 0\big\} \;. \tag{1.5}$$

For simplicity, assume that β is finite and that the supremum is attained. Then

$$\mathbb{E}[Xe^{\beta X}] \quad = \quad 0 \;.$$

As it turns out, there are three different regimes, depending on β. Namely, we distinguish the ***weakly subcritical*** ($0 < \beta < 1$), the ***intermediately subcritical*** ($\beta = 1$), and the ***strongly subcritical*** case ($\beta > 1$). They are characterized by different asymptotics of $\mathbb{P}(Z_n > 0)$ and different limit behavior, conditioned on survival. New limit theorems for the intermediately subcritical case will be proved in Chapter 3.

In the following, a major tool will be the change of measure. If we make an exponential change of measure with parameter β, we get a new measure under which S does not have any drift. More precisely, let $\varphi : \mathbb{R}^{n+1} \times \Delta^n \to \mathbb{R}$ be a bounded and measurable function. Then the probability measure $\mathbf{P}$ is defined by

$$\mathbf{E}\big[\varphi(Z_0, \dots, Z_n, Q_1, \dots, Q_n)\big] \quad := \quad e^{\gamma n}\mathbb{E}\big[\varphi(Z_0, \dots, Z_n, Q_1, \dots, Q_n)e^{-(\beta\wedge 1)(S_n - S_0)}\big] \;, \tag{1.6}$$

associated random walk	$\mathbb{P}(Z_n > 0)$	Classification
$\mathbb{E}[X] > 0$	$\mathbb{P}(Z_n > 0,\ \forall n \geq 0) > 0$ [11]	supercritical
$\mathbb{E}[X] = 0$	$\mathbb{P}(Z_n > 0) \sim \theta_1 \mathbb{P}(L_n \geq 0)$	critical
$\mathbb{E}[X] < 0$	$\mathbb{P}(Z_n > 0) \sim e^{-\gamma n + o(n)}$	subcritical

Table 1.1: Classification of BPREs. The constant $0 < \theta_1 < \infty$ depends on properties of the offspring distributions, and $0 < \gamma = -\frac{1}{n} \log \mathbb{E}[e^{L_n}]$ (see Lemma 4.3.5 in Chapter 4).

associated random walk	$\mathbb{P}(Z_n > 0) \sim$	Classification
$\mathbb{E}[Xe^X] > 0$	$\theta_2 \mathbb{P}(L_n \geq 0)$	weakly subcr.
$\mathbb{E}[Xe^X] = 0$	$\theta_3 e^{-\gamma n} \mathbf{P}(\tau_n = n)$	intermediately subcr.
$\mathbb{E}[Xe^X] < 0$	$\theta_4 \mathbb{E}[e^X]^n$	strongly subcr.

Table 1.2: Classification of subcritical BPREs ($\mathbb{E}[X] < 0$). The positive constants θ_2, θ_3 and θ_4 depend on properties of the offspring distributions, and $0 < \gamma = -\frac{1}{n} \log \mathbb{E}[e^{L_n}]$ (see Lemma 4.3.5 in Chapter 4).

with

$$\gamma \quad := \quad -\log \mathbb{E}\big[e^{(\beta \wedge 1)X}\big] \ . \tag{1.7}$$

In the weakly and intermediately subcritical cases, (1.5) translates to $\mathbf{E}[X] = 0$ and S becomes a recurrent random walk under $\mathbf{P}$. The change of measure then allows us to use many important properties of recurrent random walks. In the strongly subcritical case ($\beta > 1$), $\mathbf{E}[X] < 0$, thus S has also a negative drift with respect to $\mathbf{P}$. There, it is not suitable to change to a measure under which S is recurrent. This will become clear later. We only give the heuristics here (see [BGK05]) and assume that (1.3) gives the right order of (up to a constant) decay of the survival probability. Then the change of measure yields (with some positive constant c)

$$\mathbb{P}(Z_n > 0) \sim c \ \ e^{-\gamma n} \ \mathbf{E}[e^{L_n - (\beta \wedge 1) S_n}] \ .$$

There will only be a considerable contribution to the expectation if $L_n - (\beta \wedge 1) S_n$ is close to zero. Now there are three different cases:

- $0 < \beta < 1$. Then, $L_n - \beta S_n$ will only be small if both L_n and S_n are close to zero. The probability of having such an excursion of length n is (up to a constant factor) asymptotically equal to $n^{-3/2}$ for a zero mean and finite variance random walk (see e.g. [GKV03]). Thus here, one expects

$$\mathbf{E}[e^{L_n - \beta S_n}] \quad \approx \quad n^{-3/2} \ .$$

- $\beta = 1$. Here, $L_n - \beta S_n$ will be small if L_n and S_n are close to each other which essentially is the case if the random walk has its minimum close to the end, $\{\tau_n = n\}$. By duality, $\mathbf{P}(\tau_n = n) = \mathbf{P}(M_n < 0)$. It is known for zero mean and finite varinance random walks that the latter probability is of the order $n^{-1/2}$ (see e.g. [GKV03]). Thus

$$\mathbf{E}[e^{L_n - S_n}] \quad \approx \quad n^{-1/2} \ .$$

- $\beta > 1$. In this case, $\mathbf{E}[X] < 0$. Since for a random walk with negative drift, $L_n - S_n$ will be of constant order, one expects

$$\mathbf{E}[e^{L_n - S_n}] \quad \approx \quad \text{const.}$$

In the strongly subcritical case, the probability of survival decreases with the same order as the expected generation size, $\mathbb{E}[Z_n | Z_0 = 1] = \mathbb{E}[e^{S_n}] = \mathbb{E}[e^X]^n$. This resembles the behavior of subcritical

[11] under an additional assumption on $Q(\{0\})$

Galton-Watson processes. However, BPREs are not classified according to the expected generation size $\mathbb{E}[Z_n] = \mathbb{E}[e^X]^n$. If $\mathbb{E}[e^X] > 1$, the process may be supercritical, critical or weakly subcritical, depending on the distribution of X. $\mathbb{E}[e^X] < 0$ only implies that the BPRE is subcritical.

Table 1.1 and 1.2 summarize the classification (for simplicity assume that the offspring distributions have finite variance and that $\mathbb{E}[|X|] < \infty$ and $\mathbb{E}[|X|e^X] < \infty$; the results hold true under more general conditions, but we refrain from giving details here).

The different regimes of BPREs will be characterized in the next chapter by recalling some known results. The intermediately subcritical case is studied in Chapter 3. There, new limit theorems will be proved which describe properties of an intermediately subcritical BPRE, conditioned on survival. In Chapter 4, large deviations of BPREs are analyzed, both for offspring distributions with geometrically bounded tails as well as for heavy-tailed offspring distributions. In Chapter 4.4, a short outlook on the problem of analyzing lower deviations of supercritical BPREs is provided. In Chapter 4.5, upper large deviations of Z, conditioned on the environment (the so-called ***quenched approach***), are studied. Finally, in Chapter 5 a simulation algorithm for conditioned BPREs is explained and in Chapter 6, a short outlook on current research is presented. Some technical results are proved in the appendix.

Chapter 2

Classification and known results for BPREs

In the following, for sequences (d_n) and (m_n), we use the notation $d_n \sim m_n$ if $d_n/m_n \to 1$ as $n \to \infty$. All limit theorems presented in this chapter are under the law $\mathbb{P}$ (i.e. averaged over the environment) which is what is called the ***annealed approach***. In contrast to this, under the ***quenched approach***, limit theorems ***conditioned on the environment*** Π are developed.

2.1 The supercritical case

Suppose that $\mathbb{E}[|X|] < \infty$ and $\mathbb{E}[X] > 0$, that is S has a positive drift. In the case of classical Galton-Watson processes, supercritical processes have a positive survival probability. For BPREs, the random fluctuations of the environment can cause a.s. extinction of the process even if $\mathbb{E}[X] > 0$. For the survival of the process, an additional integrability condition for the probability that an individual has no offspring, $Q(\{0\})$, is needed. The following theorem is proved in [SW69] and [Smi68].

Theorem 2.1.1. (Smith and Wilkinson (1968/69)) *Suppose* $\mathbb{E}[|X|] < \infty$. *Then the BPRE* $(Z_n)_{n\in\mathbb{N}_0}$ *has a positive survival probability,*

$$\lim_{n\to\infty} \mathbb{P}(Z_n > 0) \quad > \quad 0 \ ,$$

iff

$$\mathbb{E}[X] > 0 \quad \text{and} \quad \mathbb{E}\Big[\log\big(1 - Q(\{0\})\big)\Big] > -\infty \ . \tag{2.1}$$

The second condition in (2.1) assures that catastrophic events, meaning the probability of an individual having no offspring is very close to one, are sufficiently improbable. Heuristically, if the second condition in (2.1) is not met, the process will die out a.s. due to such 'catastrophes' within a few generations.

For a simple example when (2.1) is not fulfilled, consider the set of distributions $\mathcal{A} \subset \Delta$ with just one free parameter, only putting mass onto two points: for all $q \in \mathcal{A}$, $l := 1 - q(0) \in (0,1)$ and $q(\lceil 2/l \rceil) = l$. Thus the mean of this distribution is

$$m_q \quad = \quad \lceil 2/l \rceil \cdot l > 1 \ .$$

Now the parameter is chosen at random according to the following distribution with parameter $\alpha \in (0,1)$,

$$\mathbb{P}(L \le x) \quad = \quad c\,(-\log(x))^{-\alpha} \ ,$$

where $c > 0$ is the norming constant. Then $\mathbb{E}[X] = \mathbb{E}\big[\log\big(\lceil 2/L \rceil \cdot L\big)\big] > 0$ and

$$\mathbb{E}\Big[\log\big(1 - Q(\{0\})\big)\Big] \ = \ \mathbb{E}[\log L] \ = \ -\infty \ .$$

2.2 The critical case

The critical case is treated in several papers (see [Koz76], [Afa93], [Afa97], [GK00], [AGKV05a], [BDKV10]). The following simulation with R [12] gives a first impression of the situation in the critical case. Survival of the process is essentially realized by a 'good' environment (i.e. $\{L_n \geq 0\}$). Conditioned on survival until generation n, S behaves like a random walk conditioned on staying positive and Z exhibits supercritical growth. This will be described more in detail by the theorems presented in this section.

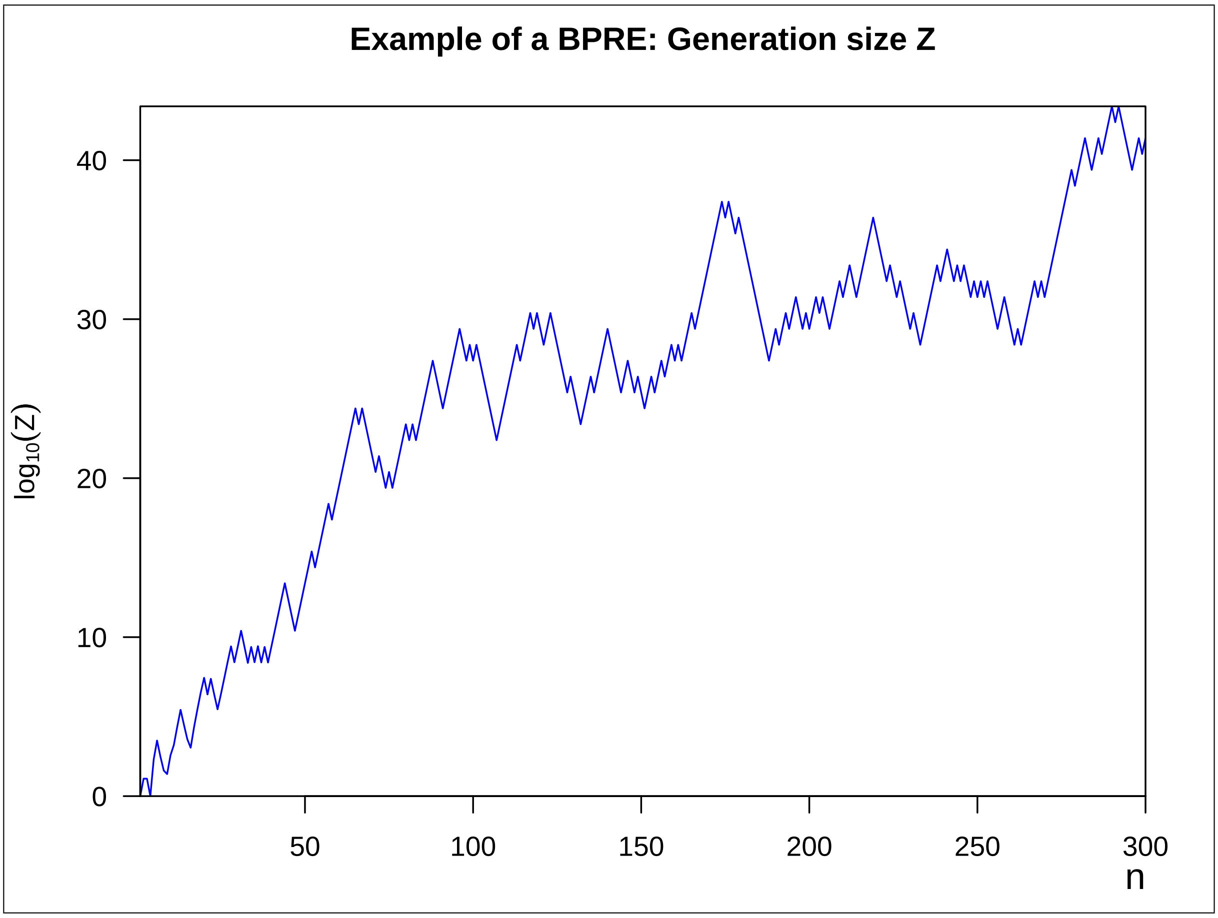


Figure 2.1: Example: A critical BPRE with geometric offspring distributions.

Here, we recall some results from [AGKV05a] on the asymptotic of the survival probability and the asymptotic behavior of the process, conditionend on survival. The following theorems are proved under the assumption that the associated random walk fulfills ***Spitzer's condition***, that is

Assumption 2.1. *There exists a number* $0 < \rho < 1$ *such that*

$$\frac{1}{n}\sum_{m=1}^{n} \mathbb{P}(S_m > 0) \quad \to \rho \quad \text{as } n \to \infty \ .$$

It says that the expected proportion of time that S spends within the positive real half line up to time n, converges as $n \to \infty$ to some value in $(0, 1)$. Any random walk fulfilling Assumption 2.1 is of the oscillating type.

Next we define renewal function

$$u(x) \quad := \quad \begin{cases} 1 + \sum_{k=1}^{\infty} \mathbb{P}(-S_k \leq x, M_k < 0) & , \quad \text{if } x \geq 0 \\ 0 & , \quad \text{else} \end{cases} \quad . \tag{2.2}$$

[12] see the R *Projekt for Statistical Computing*, www.r-project.org

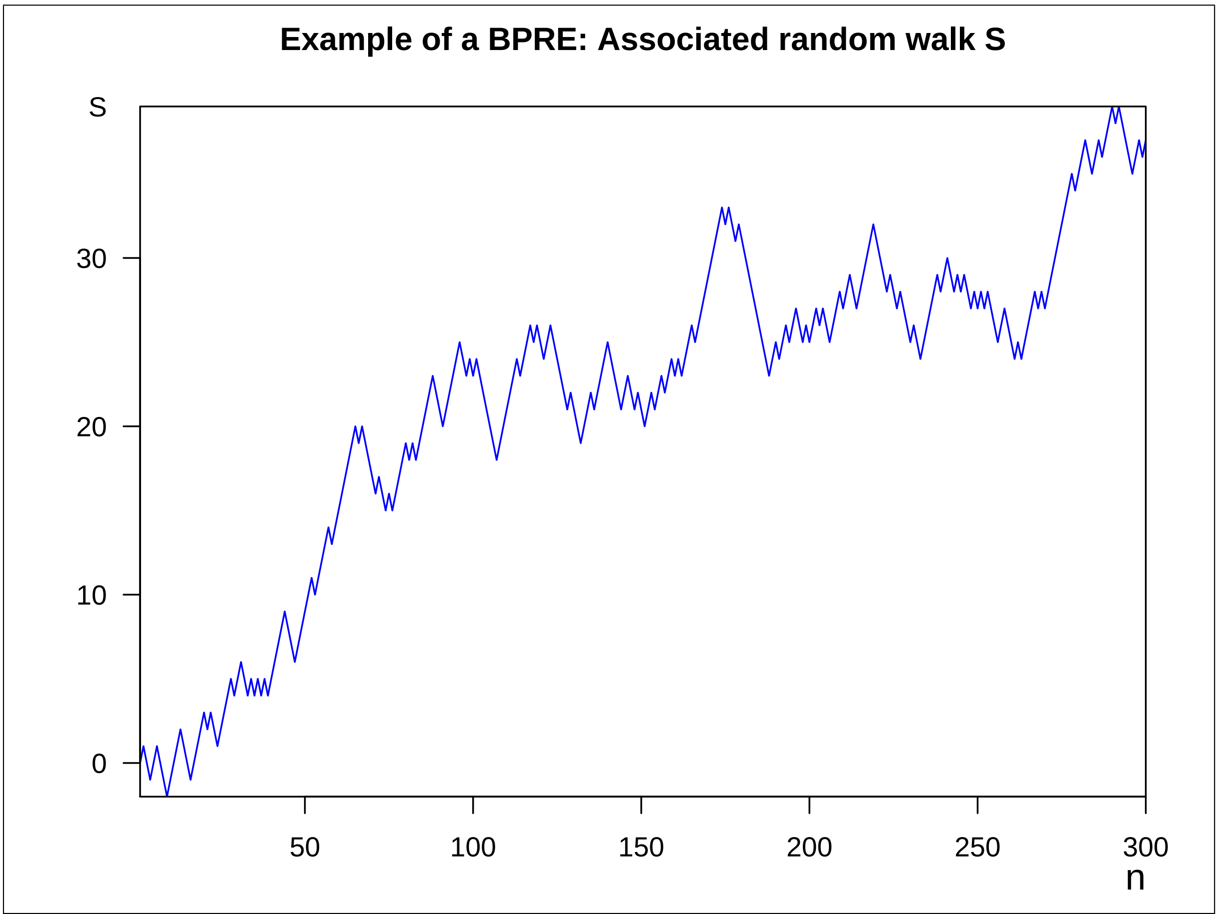


Figure 2.2: Example: Associated random walk S for the BPRE displayed in Figure 2.1.

The function u will be explained in detail later. Here just note that for any oscillating random walk,

$$\mathbb{E}[u(x+X)] \quad = \quad u(x) \quad , \; x \geq 0 \; . \tag{2.3}$$

Furthermore, one needs some regularity of the offspring distributions Q. For this, we introduce the ***standardized truncated second moment*** of Q,

$$\zeta(a) \quad := \quad \sum_{y=a}^{\infty} y^2 Q(\{y\})/m(Q)^2 \; , \quad a \in \mathbb{N}_0 \; , \tag{2.4}$$

where

$$m(Q) \; := \; \sum_{k=0}^{\infty} k \; Q(\{k\}) \qquad .$$

The second assumption now assures some regularity of the offspring distributions.

Assumption 2.2. *For some $\epsilon > 0$ and some $a \in \mathbb{N}_0$, we assume*

$$\mathbb{E}\Big[\big(\log^+ \zeta(a)\big)^{1/\rho+\epsilon}\Big] \; < \; \infty \quad \textit{and} \quad \mathbb{E}\Big[u(X)\big(\log^+ \zeta(a)\big)^{1+\epsilon}\Big] \; < \; \infty \; , \tag{2.5}$$

where $\log^+ x := \log(\max(x,1))$*.*

Some examples where Assumption 2.2 is fulfilled are:

- Q has uniformly bounded support, i.e. there exists a $c < \infty$ such that $Q(\{0,1,\ldots,c\}) = 1$ $\mathbb{P}$–a.s. In particular, Assumption 2.2 is trivially fulfilled for any binary branching process in random environment, that is when an individual has either two children ore none.

- By (2.3), $\mathbb{E}[u(X)] = u(0) < \infty$, thus Assumption 2.2 is fulfilled if $\zeta(a)$ is a.s. bounded from above. In case of Poisson or geometric distributions, the standardized second factorial moment,

$$\eta \quad := \quad \sum_{y=0}^{\infty} y(y-1)Q(\{y\})/m(Q)^2 \ ,$$

  is a constant ($\eta = 1$ for a Poisson distribution, $\eta = 2$ for a geometric distribution). As

$$\zeta(2)/2 \quad \leq \quad \eta \ ,$$

  Assumption 2.2 is fulfilled if Q is a.s. a Poisson or a geometric distribution.

- It is possible to get rid of the renewal function u in Assumption 2.2. It is known that $u(x) = O(x)$ as $x \to \infty$ (see [Fel87], Chapter XII) and by definition, $u(x) = 0$ for $x < 0$. Thus, using Hölder's inequality, Assumption 2.2 is fulfilled if

$$\mathbb{E}[(X^+)^p] < \infty \quad \text{and} \quad \mathbb{E}\big[(\log^+ \zeta(a))^q\big] < \infty$$

  for some $p > 1$ and $q > \max\{\rho^{-1}, p/(p-1)\}$.

Alternatively, if one has more regularity of the tails of X, one can replace the assumptions by the following two conditions.

Assumption 2.3. *The distribution of X has finite variance or (more generally) belongs to the domain of attraction of some stable law $s(\cdot)$ with index $\alpha \in (0, 2]$. The limit law is not a one-sided stable law, that is, $0 < s(\mathbb{R}^+) < 1$.*

This means that there is an increasing sequence of positive numbers

$$a_n \quad = \quad n^{1/\alpha} l_n$$

with a slowly varying sequence $l_1, l_2, \ldots$ such that for $n \to \infty$

$$\mathbb{P}(S_n/a_n \in dx) \quad \to \quad s(x)dx \ .$$

Remark. *In general, Assumption 2.1 is less restrictive than Assumption 2.3. However, if X^- has finite second moment, Assumption 2.3 is equivalent to Spitzer's condition (cf. [Don77]).*

By this gain of regularity, Assumption 2.2 may be relaxed to

Assumption 2.4. *For some $\epsilon > 0$ and some $a \in \mathbb{N}_0$, let*

$$\mathbb{E}\Big[\big(\log^+ \zeta(a)\big)^{\alpha+\epsilon}\Big] \quad < \quad \infty \ ,$$

where $\log^+ x := \log(\max(x, 1))$.

The following theorems have been proved in [AGKV05a]. The first one describes the assymptotic of the nonextinction probability.

Theorem 2.2.1. (Afanasyev, Geiger, Kersting and Vatutin (2005)) *Assume Assumptions 2.1 and 2.2 or 2.3 and 2.4. Then there exists a number $0 < \theta < \infty$ such that*

$$\mathbb{P}(Z_n > 0) \quad \sim \quad \theta \ \mathbb{P}(\min\{S_1, \ldots, S_n\} \geq 0) \quad \textit{as } n \to \infty \ .$$

The asymptotic of the nonextinction probability is again – up to the constant factor θ – completely determined by properties of the associated random walk. The theorem essentially says that survival until time n is -up to the constant factor- as improbable as the minimum of the associated random walk being nonnegative. This reflects the fact that by (1.3), the nonextinction probability is small if S has a low minimum. Conditioning on survival, that is on the event $\{Z_n > 0\}$, is essentially the same as conditioning on $\{\min\{S_1, \ldots, S_n\} \geq 0\}$. A more detailed description of this phenomenon is provided by Theorems 2.2.4 and 2.2.5.

Under Spitzer's condition (Assumption 2.1), the asymptotic behavior of the minimum of a random walk is well-known, which leads to the following corollary.

Corollary 2.2.2. (Afanasyev, Geiger, Kersting and Vatutin (2005)) *Assume Assumptions 2.1 and 2.2 or 2.3 and 2.4. Then there is a slowly varying sequence* $l_1, l_2, \ldots$ *such that*

$$\mathbb{P}(Z_n > 0) \quad \sim \quad \theta \; l_n \; n^{-(1-\rho)} \qquad , \; as \; n \to \infty \; .$$

Let us now look at the behavior of the branching process, conditioned on survival. As it turns out, conditioned on $\{Z_n > 0\}$, the generation size process $Z_0, Z_1, Z_2, \ldots$ shows a kind of 'supercritical' behavior. For classical Galton-Watson processes, this means that Z_n/e^{S_n} converges a.s. to some typically nondegenerated, positive random variable. In our case, the conditional distribution of the environment Π, given $\{Z_n > 0\}$, changes with n.
Thus, the following theorem is formulated for the rescaled generation size process $X^{r,n} = (X_t^{r,n})_{0\le t\le 1}$, with $r \le n$, $r, n \in \mathbb{N}_0$. It is defined by

$$X_t^{r,n} \quad := \quad \frac{Z_{r+\lfloor (n-r)t \rfloor}}{e^{S_{r+\lfloor (n-r)t \rfloor}}} \; , \quad 0 \le t \le 1 \; . \tag{2.6}$$

Theorem 2.2.3. (Afanasyev, Geiger, Kersting and Vatutin (2005)) *Assume Assumptions 2.1 and 2.2 or 2.3 and 2.4. Let* $r_1, r_2, \ldots$ *be a sequence of natural numbers such that* $r_n \le n$ *and* $r_n \to \infty$*. Then*

$$\mathcal{L}(X^{r_n,n} | Z_n > 0) \quad \overset{d}{\to} \quad \mathcal{L}((W_t)_{0\le t\le 1}) \quad as \; n \to \infty \; ,$$

where the limiting process is a stochastic process with a.s. constant paths, that is, $\mathbb{P}(W_t = W \; for \; all \; t \in [0,1]) = 1$ *for some random variable* W*. Furthermore,*

$$\mathbb{P}(0 < W < \infty) \quad = \quad 1 \; .$$

By $\overset{d}{\to}$, we denote the weak convergence with respect to the Skorohod topology in the space $\mathcal{D}[0,1]$ of càdlàg functions[13] on the unit interval. By Theorem 2.2.3, again the growth of Z is mainly determined by the associated random walk, that is by the sequence $(e^{S_n})_{n>0}$. The process thus exhibits supercritical behavior. The fine structure of the environment only affects the random variable W.

As mentioned above, conditioning on $\{Z_n > 0\}$ affects the environment and thus changes the behavior of S. The following two theorems describe this phenomenon. Recall that by τ_n, we denote the first time, when the minimum of $S_0, \ldots, S_n$ is attained (see Definition (1.4)),

$$\tau_n \quad := \quad \min \big\{ i \le n | S_i = \min\{S_0, \ldots, S_n\} \big\} \; , \quad n \ge 0 \; .$$

Theorem 2.2.4. (Afanasyev, Geiger, Kersting and Vatutin (2005)) *Assume Assumptions 2.1 and 2.2 or 2.3 and 2.4. Then, as* $n \to \infty$*,*

$$\mathcal{L}\big((\tau_n, \min\{S_0, \ldots, S_n\}) | Z_n > 0\big)$$

converges weakly to some probability measure on $\mathbb{N}_0 \times \mathbb{R}_0^-$*.*

This theorem states that, conditioned on survival, the associated random walk has its (global) minimum at some finite time.
A more detailed description is proved in the situation of Assumption 2.3.

[13] continue à droite, limite à gauche

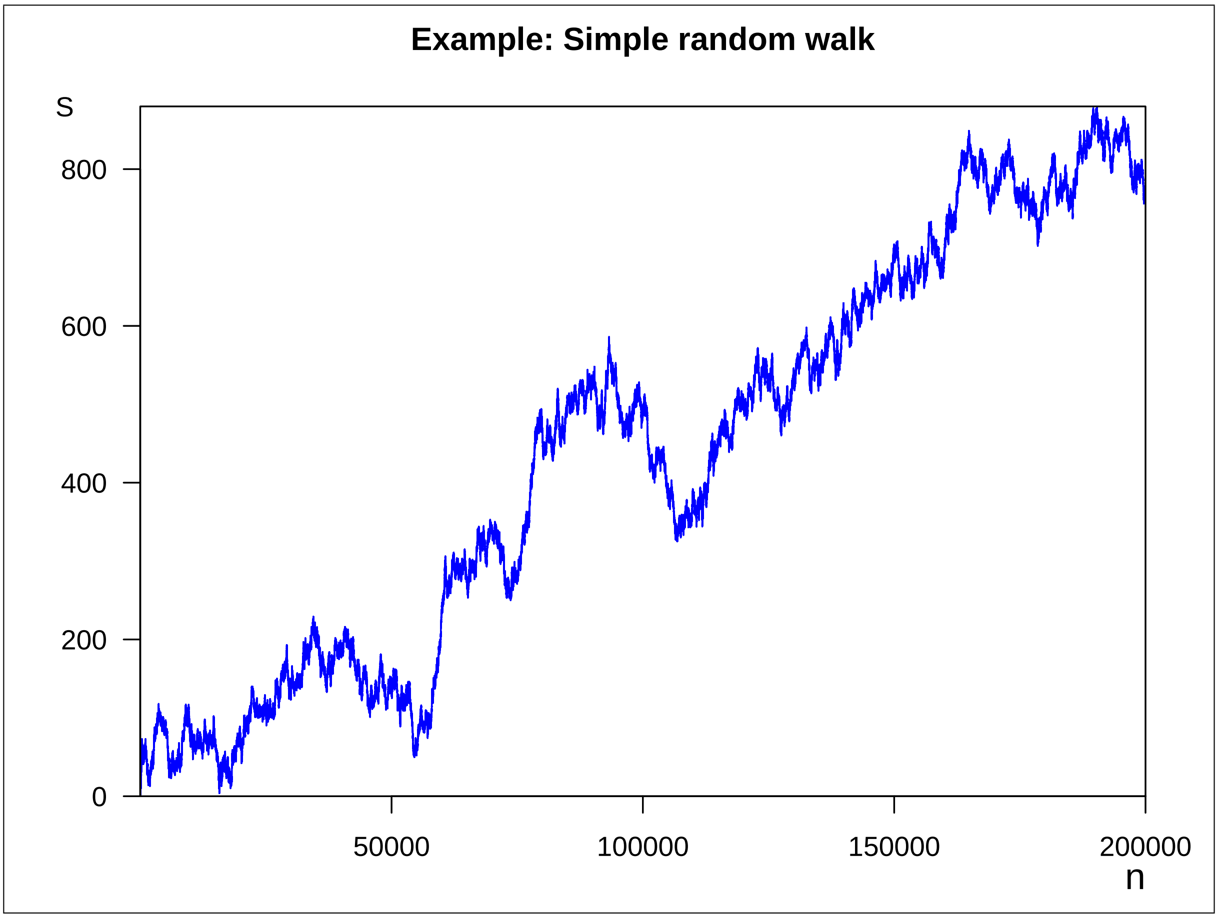


Figure 2.3: Example: A simple random walk conditioned to stay nonnegative.

Theorem 2.2.5. (Afanasyev, Geiger, Kersting and Vatutin (2005)) *Assume Assumptions 2.3 and 2.4. Then there exists a slowly varying sequence $l_1, l_2, \ldots$ such that*

$$\mathcal{L}((n^{-1/\alpha} l_n S_{\lfloor nt \rfloor})_{0 \le t \le 1} | Z_n > 0) \quad \xrightarrow{d} \quad \mathcal{L}(L^+) \quad \text{as } n \to \infty \ ,$$

where L^+ denotes the meander of a strictly stable Lévy process L with index α.

Let us shortly explain the process L^+. Convergence of a conditioned Brownian motion to the Brownian meander has been proved in [DIM77] (see also e.g. [Don85] and [Dur78] for convergence of conditioned Markov chains). Essentially, L^+ is a strictly stable Lévy process, conditioned to stay positive on the time intervall $(0, 1]$. L^+ is the limiting process of $\{(n^{-1/\alpha} l_n S_{\lfloor nt \rfloor})_{0 \le t \le 1} | L_n \ge 0\}$. The existence and characterization of this limit can be found in [Don85]. Figure 2.3 illustrates a random walk conditioned to stay nonnegative. In case of finite variance, $\alpha = 2$, L^+ is the meander of a standard Brownian motion. To sum up, a critical BPRE behaves, conditioned on survival, similar to a supercritical branching process. Conditioned on survival, the environment is 'good', i.e. the rescaled associated random walk converges to Lévy process conditioned to stay nonnegative.

2.3 The subcritical cases

2.3.1 The strongly subcritical case

This case has been discussed intensively in [AGKV05a]. Here we recall the main results. The existence of β (defined in 1.5) is actually not needed. It suffices that the following condition is fulfilled:

Assumption 2.5.

$$\mathbb{E}[Xe^X] \quad < \quad 0 \ .$$

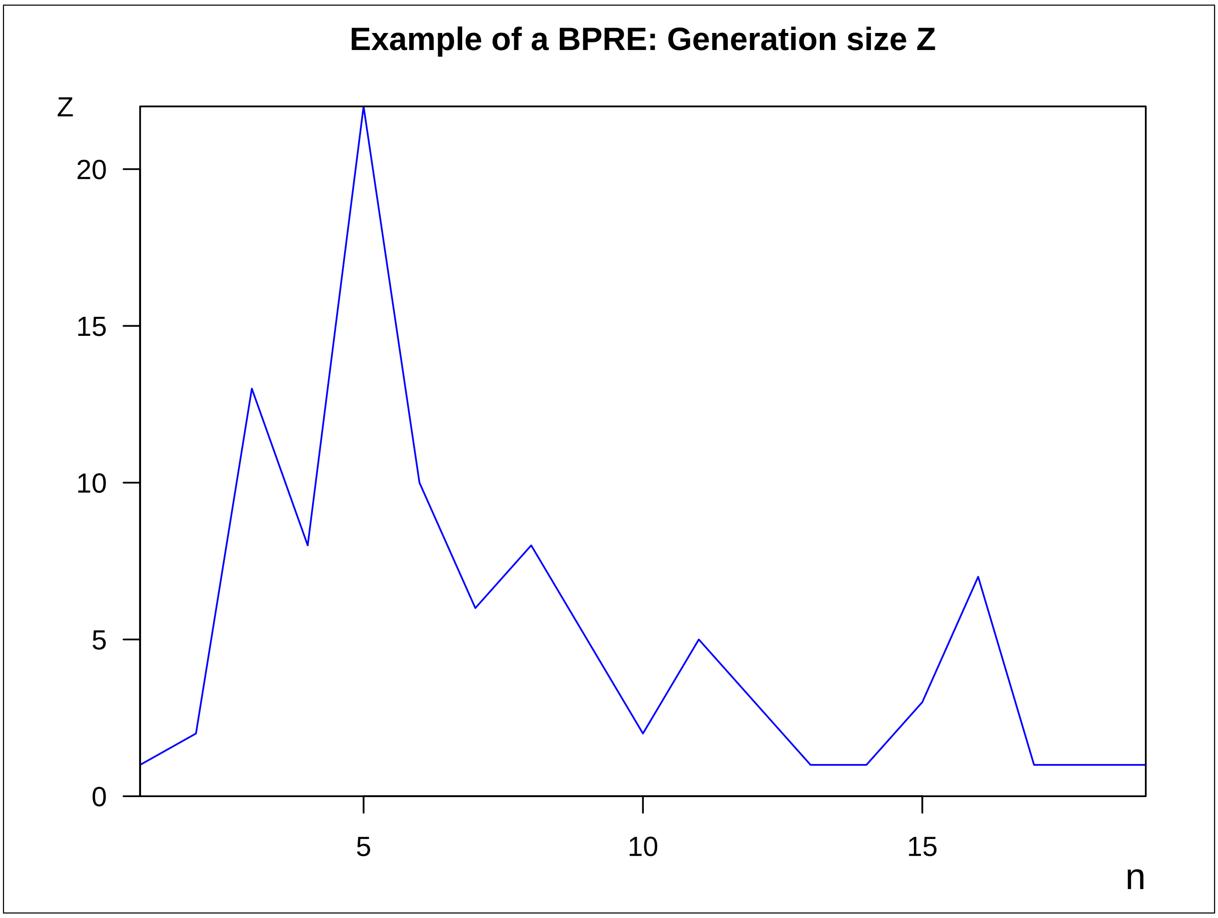


Figure 2.4: Example: A strongly subcritical BPRE with geometric offspring distributions

See Figures 2.4 and 2.5 for sample simulations with R. Additionally, an integrability condition for the offspring distributions is required, namely:

Assumption 2.6.

$$\mathbb{E}[Z_1 \log^+ Z_1] \quad < \quad \infty \ ,$$

where $\log^+ x := \log(\max(x,1))$.

Note that Assumption 2.6 implies $\mathbb{E}[Z_1] = \mathbb{E}[m(Q)] < 1$ and $\mathbb{E}[\log Z_1] = \mathbb{E}[\log m(Q)] < 0$. An assumption on the standardized second factorial moment of Z_1, conditioned on Π (see Definition (2.4)) with Assumption 2.5 already implies Assumption 2.6. Let for some $a > 0$,

$$\mathbb{E}[m(Q) \log^+ \zeta(a)] \quad < \quad \infty \ . \tag{2.7}$$

Then by Jensen's inequality and the definition of ζ (see (2.4)),

$$\begin{aligned} \sum_{y=1}^{\infty} \log y \Big(\frac{yQ(\{y\})}{m(Q)} \Big) \ &\leq \ a \log a + \log \Big(\sum_{k=a}^{\infty} \frac{y^2 Q(\{y\})}{m(Q)} \Big) \\ &\leq \ a \log a + \log^+ m(Q) + \log^+ \zeta(a) \quad \mathbb{P}-\text{a.s.} \end{aligned}$$

Multiplying both sides with $m(Q)$ and taking the expectation yields

$$\mathbb{E}[Z_1 \log^+ Z_1] \quad \leq \quad a \log a \mathbb{E}[m(Q)] + \mathbb{E}[m(Q) \log^+ m(Q)] + \mathbb{E}[m(Q) \log^+ \zeta(a)] \ < \ \infty \ .$$

For examples where (2.7) is fulfilled, we refer to Section 2.2. In particular, (2.7) is abundant if Q is $\mathbb{P}$–a.s. a geometric or a Poisson distribution. Also note that taking $\mathbb{P}(Q = q) = 1$ for some $q \in \Delta$ yields a

Figure 2.5: Example: Associated random walk S.

classical Galton-Watson process. If $m(q) < 1$, Assumption 2.5 is fulfilled. The second assumption is then well-known (see e.g. [AN72, p. 45]) to be a sufficient condition for

$$\mathbb{P}(Z_n > 0) \quad \sim \quad c\,\mathbb{E}[Z_n] \ = \ c\,m(q)^n \ .$$

Like in the case of classical Galton-Watson processes, the first moment estimate already gives the right decay rate of the survival probability in the strongly subcritical regime (see [GL01], originally proved in [DH97b] under an additional moment condition).

Theorem 2.3.1. (Guivarc'h and Liu (2001)) *Under Assumptions 2.5 and 2.6, there is a $0 < \theta \leq 1$ such that*

$$\mathbb{P}(Z_n > 0) \quad \sim \quad \theta\,\mathbb{E}[Z_n] \quad \text{as } n \to \infty \ .$$

The next theorem, proved in [GKV03], states that – as in the case of subcritical Galton-Watson processes – the generation size has a weak limit, conditioned on nonextinction.

Theorem 2.3.2. (Geiger, Kersting and Vatutin (2003)) *Under Assumptions 2.5 and 2.6, there is a probability measure υ with weights υ_z, such that*

$$\lim_{n\to\infty} \mathbb{P}(Z_n = z | Z_n > 0) \quad = \quad \upsilon_z, \quad z \in \mathbb{N} \ .$$

The previous theorem, together with Fatou's lemma, yields $m(\upsilon) \leq \theta^{-1} < \infty$ so that the expectation with respect to υ is finite. In fact, as it is proved in [AGKV05a], $m(\upsilon) = \theta^{-1}$. The next theorems describe the behavior of a strongly subcritical branching process in random environment, conditioned on survival, more in detail. The following theorem says that, conditioned on nonextinction, the offspring distributions are independent in the limit with respect to $\mathbf{P}$ (recall definition (1.6)). Moreover, excursions of the associated random walk vanish in the skaling limit. This means that the process does not exhibit any supercritical behavior, conditioned on survival.

Theorem 2.3.3. (Afanasyev, Geiger, Kersting and Vatutin (2005)) *Assume 2.5 and 2.6, let $i_{n,j}, n \in \mathbb{N}, 1 \le j \le k$ be nonnegative integers with $1 \le i_{n,1} < i_{n,2} < \cdots < i_{n,k} \le n$, and $n - i_{n,k} \to \infty$ as $n \to \infty$. Then for every $k \in \mathbb{N}$ and Borel sets $B_1, \ldots, B_k \subset \Delta$,*

$$\lim_{n\to\infty} \mathbb{P}(Q_{i_{n,1}} \in B_1, \ldots, Q_{i_{n,k}} \in B_k | Z_n > 0) \quad = \quad \prod_{j=1}^{k} \mathbf{P}(Q \in B_j) \ .$$

Moreover,

$$\mathcal{L}\big((n^{-1} S_{\lfloor nt \rfloor})_{0\le t\le 1} \big| Z_n > 0\big) \quad \stackrel{d}{\to} \quad \mathcal{L}\big((t\mathbf{E}[X])_{0\le t\le 1}\big)$$

with respect to the Skorohod topology.

The next result characterizes the dynamics of the generation size process Z, conditioned on nonextinction. No scaling of Z is necessary, which essentially means that, conditioned on $\{Z_n > 0\}$, the population stays small throughout the time intervall from 0 to n. In the special case when Q has linear fractional generating functions, the following theorem has first been obtained in [Afa01b]. A more general version of it has been established in [AGKV05a].

Theorem 2.3.4. (Afanasyev, Geiger, Kersting and Vatutin (2005)) *Under Assumptions 2.5 and 2.6 and for any $0 < t_1 < \cdots < t_k < 1$, as $n \to \infty$*

$$\mathcal{L}\big((Z_{\lfloor nt_1 \rfloor}, \ldots, Z_{\lfloor nt_k \rfloor})\big| Z_n > 0\big) \quad \stackrel{d}{\to} \quad (W_1, \ldots, W_k) \ ,$$

where $W_1, W_2, \ldots$ are i.i.d. copies of some random variable W with

$$\mathbb{P}(1 \le W < \infty) \quad = \quad 1 \ .$$

Summing up, survival of the process in the strongly subcritical regime is typically not realized by a favorable environment (i.e. $\min\{S_1, \ldots, S_n\} \ge 0$ or bounded), although the environment is more favorable than it is typically for the unconditioned process (i.e. $\mathbf{E}[X] > \mathbb{E}[X]$). In the limit and conditioned on nonextinction, the associated random walk behaves in a completely deterministic manner and decays linearly according to its expectation with respect to the measure $\mathbf{P}$ and the offspring distributions at different times become independent. Survival is realized by exceptional offspring numbers within this environment.

2.3.2 The weakly subcritical case

The results in this section are based on joint work with Valery Afanasyev[14], Götz Kersting and Vladimir Vatutin[14] and published in [ABKV10]. As it turns out, methods developed in [AGKV05a] for criticality, can also be used for weak subcriticality. Conditioned on survival, the process exhibits 'supercritical' behavior (see Figures 2.6 and 2.7 for simulations).

Let us briefly state the results from [ABKV10]:

Assumption 2.7. *The process Z is weakly subcritical, that is there is a number $0 < \beta < 1$ such that*

$$\mathbb{E}[Xe^{\beta X}] \ = \ 0 \ .$$

As explained in Section 1.2, this allows to change to a measure $\mathbf{P}$ according to (1.6) and S becomes a recurrent random walk under $\mathbf{P}$.

As to the regularity of the distribution of X the following assumptions are needed.

Assumption 2.8. *The distribution of X has finite variance with respect to $\mathbf{P}$ or (more generally) belongs to the domain of attraction of some stable law $s(\cdot)$ with index $\alpha \in (1, 2]$. It is non-lattice.*

[14]Steklov Institute, Moscow

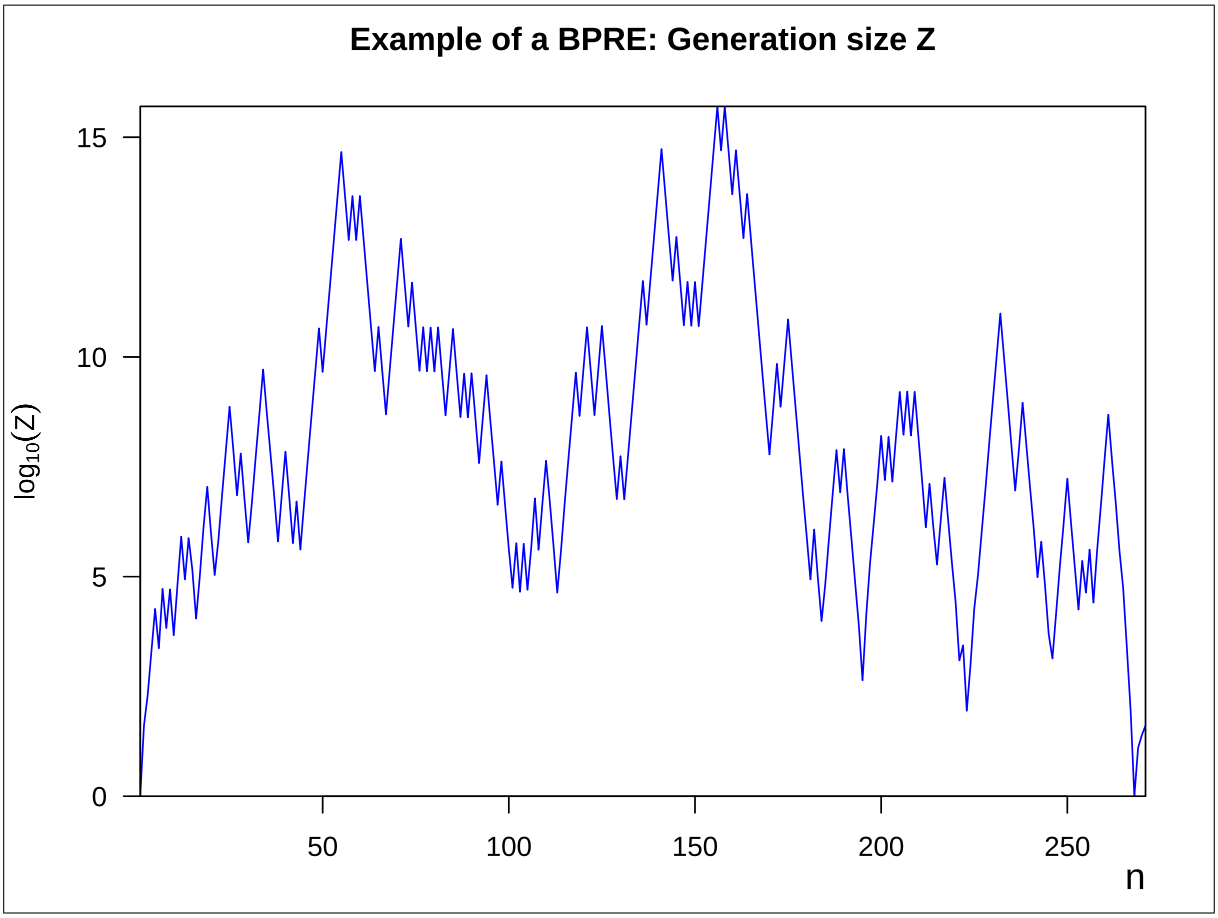


Figure 2.6: Example: Logarithm of the population size of a weakly subcritical BPRE with geometric offspring distributions.

The last assumption on the environment concerns the standardized truncated second moment of Q. Recall definition (2.4),

$$\zeta(a) \;:=\; \sum_{y=a}^{\infty} y^2 Q(\{y\}) \Big/ m(Q)^2 \;, \quad a \in \mathbb{N} \;.$$

Assumption 2.9. *For some $\varepsilon > 0$ and some $a \in \mathbb{N}$*

$$\mathbf{E}[(\log^+ \zeta(a))^{\alpha+\varepsilon}] \;<\; \infty \;,$$

where $\log^+ x := \log(\max(x,1))$.

Compare Section 2.2 for examples where this assumption is fulfilled.

Now the main results for the weakly subcritical case will be described briefly. The first theorem describes the asymptotic behavior of the nonextinction probability at generation n.

Theorem 2.3.5. (Afanasyev, B., Kersting and Vatutin (2009)) *Under Assumptions 2.7 to 2.9, there exists a number $0 < \kappa < \infty$ such that*

$$\mathbb{P}(Z_n > 0) \;\sim\; \kappa\, \mathbb{P}(\min(S_1, \ldots, S_n) \geq 0) \quad \textit{as } n \to \infty \;.$$

Recall that, as it has been mentioned in Section 2.2, under Assumption 2.8, there exists an increasing sequence of positive numbers

$$a_n \;=\; n^{1/\alpha} l_n \;,$$

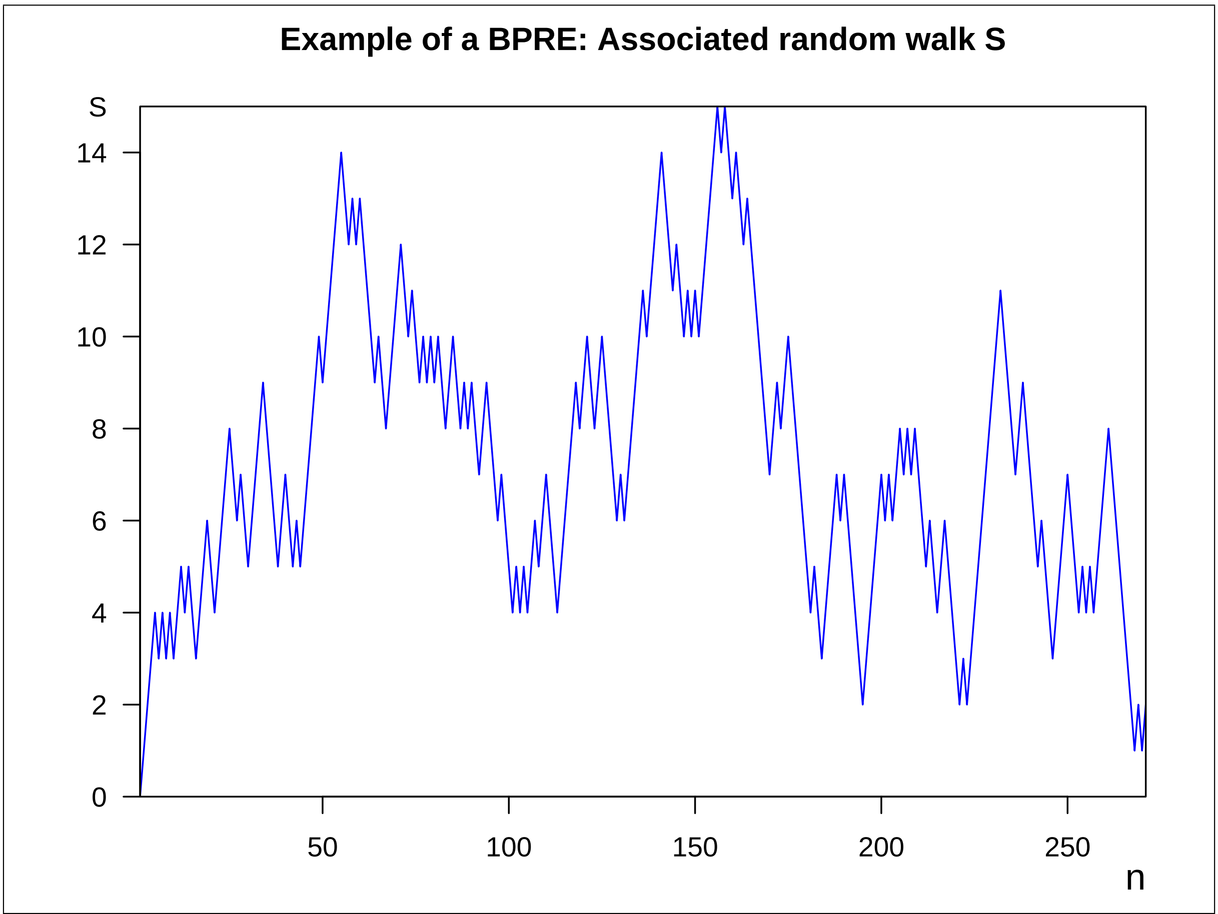


Figure 2.7: Example: Associated random walk S.

with a slowly varying sequence $l_1, l_2, \ldots$ such that for $n \to \infty$,

$$\mathbf{P}(S_n/a_n \in dx) \quad \to \quad s(x)dx \ .$$

Then from Theorem 2.3.5, the following corollary results:

Corollary 2.3.6. *Under Assumptions 2.7 to 2.9, there is a number $0 < \kappa' < \infty$ such that*

$$\mathbb{P}(Z_n > 0) \ \sim \ \kappa' \frac{e^{-\gamma n}}{n \ a_n} \ ,$$

where $\gamma = -\log \mathbb{E}[e^{\beta X}]$.

The next theorem yields convergence of the laws of $(Z_n)_{n\in\mathbb{N}}$, conditioned on survival.

Theorem 2.3.7. (Afanasyev, B., Kersting and Vatutin (2009)) *Under Assumptions 2.7 to 2.9, the conditional laws $\mathcal{L}(Z_n \,|\, Z_n > 0)$, $n \geq 1$, converge weakly to some probability distribution on the natural numbers. Moreover the sequence $\mathbb{E}[Z_n^\vartheta \,|\, Z_n > 0]$ is bounded for any $\vartheta < \beta$, implying convergence to the corresponding moment of the limit distribution.*

The last theorem describes the limiting behavior of the rescaled generation size process $e^{-S_k}Z_k$ for $r_n \leq k \leq n - r_n$, where (r_n) is a sequence of natural numbers with $r_n \to \infty$ (and $r_n < n/2$). Thus we consider the process $Y^n = \{Y_t^n, t \in [0,1]\}$, defined by

$$Y_t^n := \exp(-S_{r_n+\lfloor (n-2r_n)t\rfloor})Z_{r_n+\lfloor (n-2r_n)t\rfloor} \ .$$

This process has asymptotic paths of a constant random value. More precisely, the following statement holds:

Theorem 2.3.8. (Afanasyev, B., Kersting and Vatutin (2009)) *Under Assumptions 2.7 to 2.9, there is a process* $\{W_t, t \in [0,1]\}$ *such that as* $n \to \infty$

$$\mathcal{L}\big(Y_t^n, t \in [0,1] \bigm| Z_n > 0\big) \stackrel{d}{\to} \mathcal{L}\big(W_t,\ t \in [0,1]\big)$$

weakly in the Skorohod space $\mathcal{D}[0,1]$. *Moreover, there is a random variable* W *such that* $W_t = W$ *a.s. for all* $t \in [0,1]$ *and*

$$\mathbb{P}\{0 < W < \infty\} \ = \ 1\ .$$

Weaker versions of these results can be found in [Afa98] and [GKV03].

Thus we have the following scenario in the weakly subcritical case: Conditioned on $\{Z_n > 0\}$, Z starts growing in a favorable environment at the beginning and roughly up to time $\lfloor \epsilon n \rfloor$ Z exhibits supercritical growth. It then follows the value of $e^{S_k} = \mathbb{E}[Z_k \,|\, \Pi]$ in a completely deterministic manner, up to a random factor $W > 0$. It is due to the fact that S_k takes large values there. This behavior persists as long as S takes large values. Close to n, in the manner of an excursion (see [DIM77]) S returns to zero and the environment becomes more and more unfavorable. The values of Z are decreasing so that Z is again small close to generation n.

Chapter 3

The intermediately subcritical case

3.1 Introduction and main results

Here we study the intermediately subcritical case (see Figures 3.1 and 3.2 for simulations). As it already can be seen from the sample simulation, in contrast to the other cases, the population stays small throughout most of the time, but still, there are supercritical periods when the population grows very large. In this chapter, this 'alternating' behavior is studied.

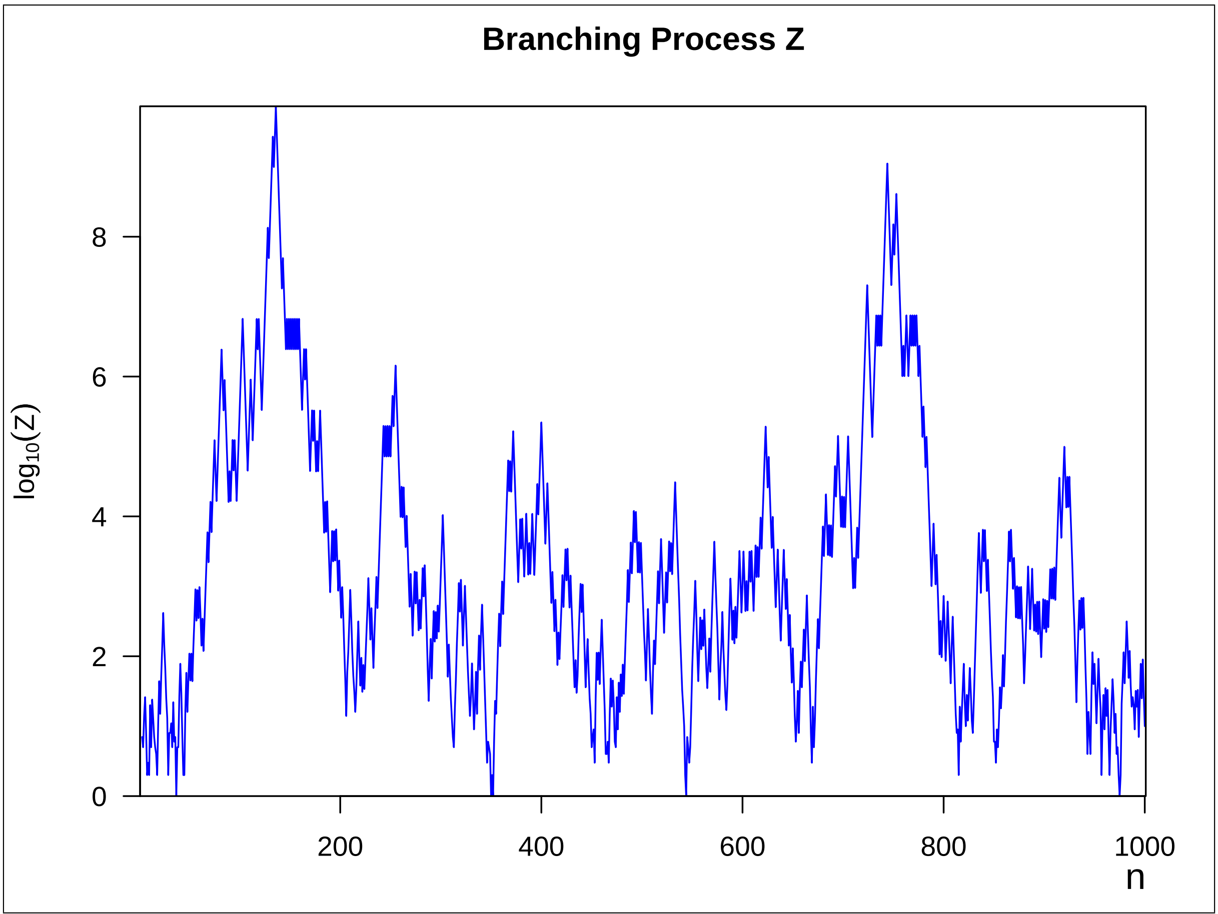


Figure 3.1: Example: An intermediately BPRE Z conditioned on $\{Z_{1000} > 0\}$ (based on the simulation scheme described in Chapter 5).

Assumption 3.1.

$$\mathbb{E}[Xe^X] \quad = \quad 0\ . \tag{3.1}$$

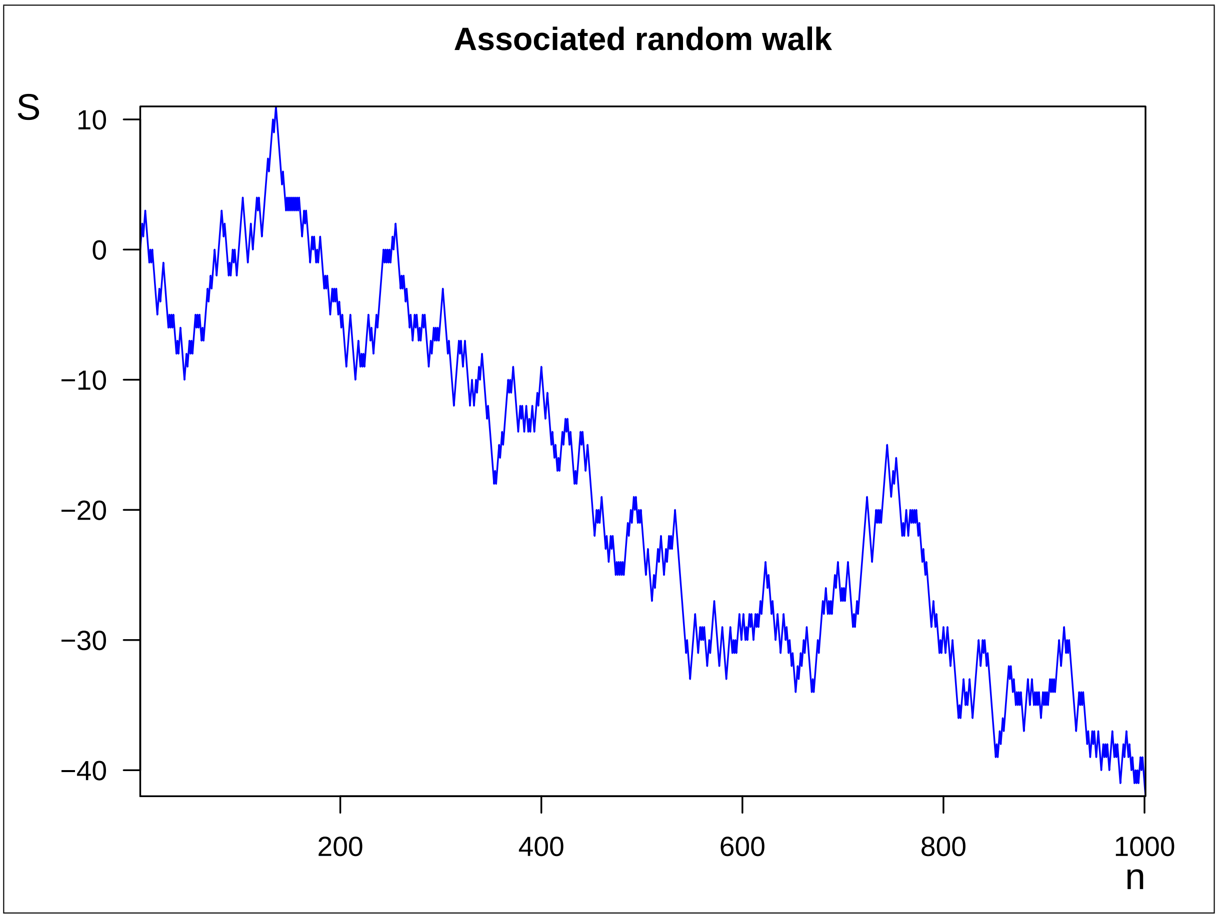


Figure 3.2: Example: Associated random walk S for the process in Figure 3.1 (simple random conditioned on having its minumum at the end).

From (3.1) results $\mathbb{E}[X] < 0$ (except for the degenerated case $X = 0$ a.s.). By Jensen's inequality it follows that (see [AGKV05b])

$$\mathbb{E}[Xe^X] \;=\; \mathbb{E}\big[\log(e^X)\cdot e^X\big] \;\geq\; \log\Big(\mathbb{E}[e^X]\Big)\cdot\mathbb{E}[e^X]$$

and thus $\mathbb{E}[e^X] < 1$.

As explained in Chapter 1, (3.1) suggests to change to the measure $\mathbf{P}$ with expectation $\mathbf{E}$ defined by

$$\mathbf{E}\left[\Phi(Q_1,\ldots,Q_n,Z_0,\ldots,Z_n)\right] \;=\; \gamma^{-n}\mathbb{E}\left[\Phi(Q_1,\ldots,Q_n,Z_0,\ldots,Z_n)e^{S_n-S_0}\right]$$

with

$$\gamma \;=\; \mathbb{E}\left[e^X\right] \;<\; 1.$$

$\mathbf{E}[X] = 0$ follows from (3.1). Thus S becomes a recurrent random walk under $\mathbf{P}$. We need a regularity assumption for the distribution of X.

Assumption 3.2. *The distribution of X has, with respect to $\mathbf{P}$, finite variance or (more generally) belongs to the domain of attraction of some strictly stable law with index $\alpha \in (1,2]$. For convenience, it is non-lattice.*

We recall a few consequences of Assumption 3.2 (see e.g. [ABKV10]). There is an increasing sequence of positive numbers

$$a_n \;=\; n^{1/\alpha}l_n$$

with a slowly varying sequence $l_1, l_2, \ldots$ such that

$$\mathbf{P}(S_n/a_n \in dx) \quad \to \quad s(x)dx$$

weakly, where $s(x)$ denotes the density of the limiting stable law. Furthermore, there is a $\rho \in (0,1)$, $\rho \leq \alpha^{-1}$ such that

$$\frac{1}{n}\sum_{m=1}^{n} \mathbf{P}(S_m > 0) \quad \to \quad \rho \quad \text{as } n \to \infty \ .$$

(see e.g. [AGKV05a]). The next assumption concerns the standardized truncated second moment of Q. Recall from (2.4) that

$$\zeta(a) \quad = \quad \sum_{y=a}^{\infty} y^2 Q(\{y\}) \Big/ m(Q)^2, \quad a \in \mathbb{N} \ ,$$

where $m_Q = \sum_{y=0}^{\infty} y m_Q(\{y\})$.

Assumption 3.3. *There are constants $0 < \epsilon < \infty$ and $a \in \mathbb{N}$ such that*

$$\mathbf{E}\big[(\log^+ \zeta(a))^{\alpha+\epsilon}\big] \ < \ \infty \ .$$

For examples where this assumption is met see Chapter 2.2.

Recall that τ_n is the time of the first minimum of $(S_0, \ldots, S_n)$ (see Definition (1.4) in Chapter 1),

$$\tau_n \quad := \quad \min\big\{0 \leq k \leq n | S_k = \min\{S_0, \ldots, S_n\}\big\} \quad , \ n \geq 0 \ .$$

The following theorem is proved in [Vat04] under somewhat stronger conditions.

Theorem 3.1.1. *Under Assumptions 3.1 to 3.3,*

$$\mathbb{P}(Z_n > 0) \quad \sim \quad \gamma^n \theta \ \mathbf{P}(\tau_n = n)$$

for some $0 < \theta < \infty$.

Remark. *There is an expression of θ in terms of a sum of expectations with respect to* $\mathbf{E}$ *(see proof of Theorem 3.1.1, Section 3.4).*

Our next theorem describes the number of times − conditioned on survival − when there is only one individual left. We assume that, with positive probability, there are distributions allowing individuals to have no child or exactly one child.

Assumption 3.4.

$$\mathbb{E}\big[Q(\{1\})Q(\{0\})\big] \quad > \quad 0 \ .$$

Then the number of times when there is just one individual left is of the same order as the number of strictly descending ladder points of an oscillating random walk conditioned on $\{\tau_n = n\}$.

Theorem 3.1.2. *Under Assumptions 3.1 to 3.4, there is a slowly varying sequence $b_1, b_2, \ldots$ such that*

$$\mathbb{E}\Big[\sharp\{k | Z_k = 1\}\Big| Z_n > 0\Big] \quad = \quad \Theta(b_n \ n^{1-\rho}) \ .$$

By $x_n = \Theta(y_n)$ we denote that the sequence x_n is of the same order as y_n as $n \to \infty$. More precisely, there are constants $c_1, c_2 \in \mathbb{R}^+$ such that

$$c_1 \ \leq \ \liminf_{n\to\infty} \frac{x_n}{y_n} \quad \leq \quad \limsup_{n\to\infty} \frac{x_n}{y_n} \ \leq \ c_2 \ .$$

As it is already indicated by Theorem 3.1.1, conditioned on nonextinction, the associated random walk converges to a Lévy process conditioned on having its minimum at the end.

Theorem 3.1.3. *Under Assumptions 3.1, 3.2, and 3.3, as $n \to \infty$,*

$$\mathcal{L}\Big(\big(S_{\lfloor nt\rfloor}/a_{\lfloor nt\rfloor}\big)_{0\le t\le 1}\big|Z_n>0\Big) \quad \overset{d}{\to} \quad \mathcal{L}(L^\dagger)$$

in distribution with respect to the Skorohod metric, where $L^\dagger$ denotes a Lévy process conditioned to have its minimum at the end.

Essentially, this theorem says that for the associated random walk, conditioning on survival of the process Z is the same as conditioning on having the minumum at the end (see Figure 3.3). Thus, survival of the process is typically not realized by a 'favorable' environment, but by exceptional offspring numbers with an environment characterized by a recurrent associated random walk having its minimum close to the end.

Typically, a recurrent random walk conditioned on having its minimum close to the end may have

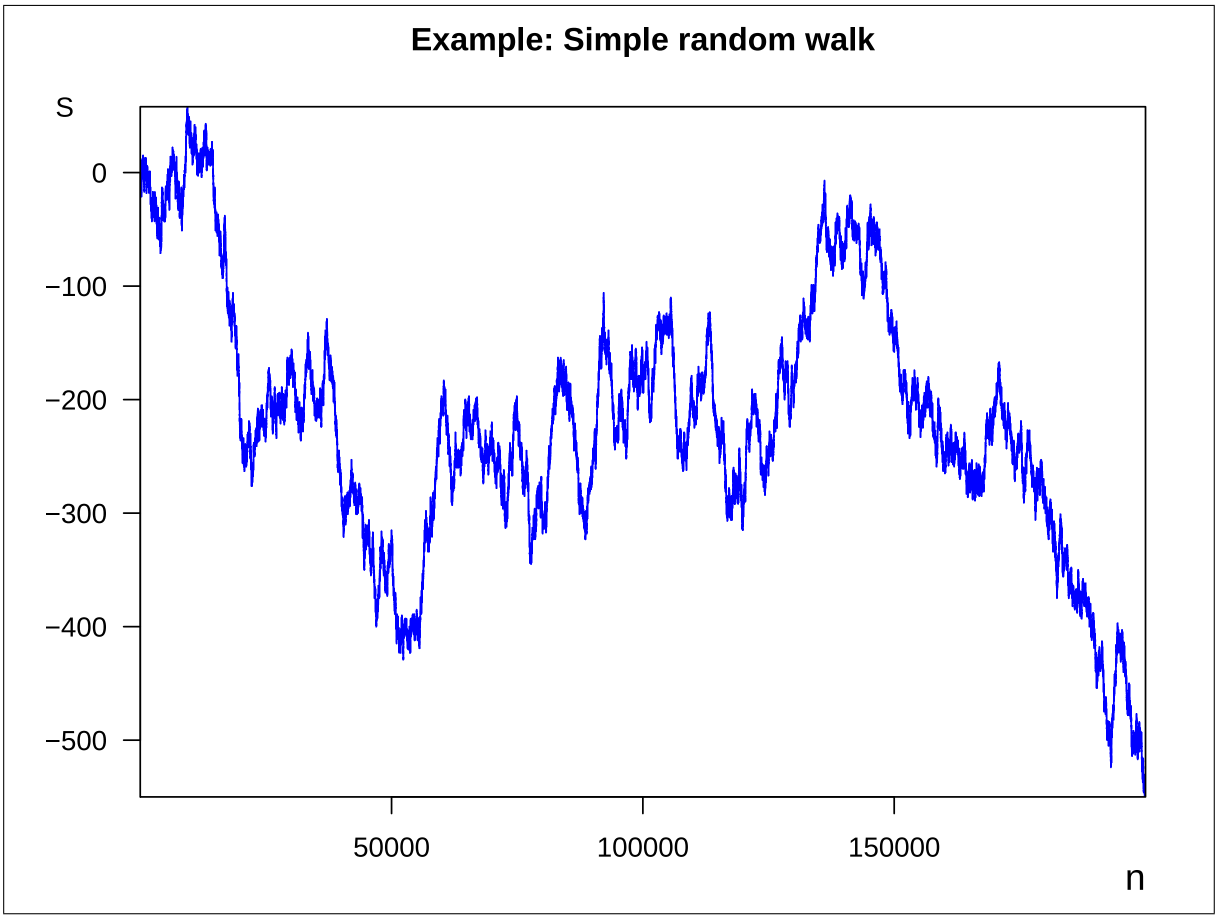


Figure 3.3: Example: A simple random walk conditioned to have its minimum at the end.

some long excursions between the strictly descending ladder points. During a long excursion, we expect that the process may grow very large and exhibits a behavior similiar to a weakly subcritical BPRE, conditioned on survival (see Section 2.3.2). For simulations of the conditioned process see Chapter 5.

A theorem describing this limiting behavior of the rescaled generation size process has been proved in [Afa01b] in the case of geometric offspring distributions. We claim that the statement also holds in the more general case treated here, but the proof remains an open problem.

Open Problem. *Let $L = (L_t^\dagger)_{0\le t\le 1}$ be a strictly stable Lévy process, conditioned to have its minimum at time* 1*. By $e_1, e_2, \dots$ we denote its excursion intervals. Let $j(t) = i$ if $t \in e_i$. Then for $0 < t_1 < t_2 < \dots < t_k < 1$, as $n \to \infty$*

$$\mathcal{L}\left(\Big(\frac{Z_{\lfloor nt_1\rfloor}}{\exp(S_{\lfloor nt_1\rfloor} - \min_{k\le\lfloor nt_1\rfloor} S_k)}, \dots, \frac{Z_{\lfloor nt_k\rfloor}}{\exp(S_{\lfloor nt_k\rfloor} - \min_{k\le\lfloor nt_k\rfloor} S_k)}\Big)\Big| Z_n > 0\right) \xrightarrow{d} \mathcal{L}\big((W_{j(t_1)}, \dots, W_{j(t_k)})\big) ,$$

where $W_1, W_2, \dots$ are i.i.d. copies of some strictly positive random variable W.

3.2 Conditional limit laws for oscillating random walks

3.2.1 A change of measure

In this section, we introduce the probability measures $\mathbf{P}^+$ and $\mathbf{P}^-$ which will be used in the proof of the limit laws (see Sections 3.3 and 3.4). Here, we state general results for oscillating random walks without referring to BPREs. Let

$$\begin{aligned} M_n &:= \max_{1\le j\le n} S_j \\ L_n &:= \min_{1\le j\le n} S_j \ . \end{aligned}$$

Let us recall some results from fluctuation theory of random walks. Two standard references are [Spi64] and [Fel87, chaper XII]. We introduce the right-continuous functions $u : \mathbb{R}_0^+ \to \mathbb{R}$ and $v : \mathbb{R}_0^- \to \mathbb{R}$ defined by

$$\begin{aligned} u(x) &:= 1 + \sum_{k=1}^\infty \mathbf{P}(-S_k \le x, M_k < 0), \quad x \ge 0 \ , \\ v(x) &:= 1 + \sum_{k=1}^\infty \mathbf{P}(-S_k > x, L_k \ge 0), \quad x \le 0 \ . \end{aligned} \tag{3.2}$$

In particular $u(0) = v(0) = 1$. Also it is known that $u(x) = O(x)$, $v(x) = O(x)$. Another representation of u and v uses the strictly descending ladder times, $0 =: \underline{\gamma}_0 < \underline{\gamma}_1 < \cdots$ and the weakly ascending ladder times, $0 =: \overline{\gamma}_0 < \overline{\gamma}_1 \cdots$, defined by

$$\begin{aligned} \underline{\gamma}_i &:= \min\left\{n > \underline{\gamma}_{i-1} \ : \ S_n < S_{\underline{\gamma}_{i-1}}\right\} \ , \\ \overline{\gamma}_i &:= \min\left\{n > \overline{\gamma}_{i-1} \ : \ S_n \ge S_{\overline{\gamma}_{i-1}}\right\} \ , \quad i \ge 1 \ . \end{aligned}$$

Then

$$\begin{aligned} u(x) &= 1 + \sum_{k=1}^\infty \mathbf{P}(S_{\underline{\gamma}_k} \ge -x), \quad x \ge 0 \ , \\ v(x) &= 1 + \sum_{k=1}^\infty \mathbf{P}(S_{\overline{\gamma}_k} < -x), \quad x \le 0 \ . \end{aligned} \tag{3.3}$$

Thus $u(x)$ is the expected number of strictly descending ladder epochs that do not fall below the level $-x$ (see Figure 3.4 for an example).

By the duality lemma, (3.2) and (3.3) are equivalent (see [Fel87, pp. 394/395]). Let us briefly recall the duality argument here for the function v as duality will also be an important tool later. In the sequel, $\hat{S}$ shall denote the dual random walk, defined by

$$\hat{S}_k \ := \ S_n - S_{n-k} \quad , \ k = 0, \dots, n \quad . \tag{3.4}$$

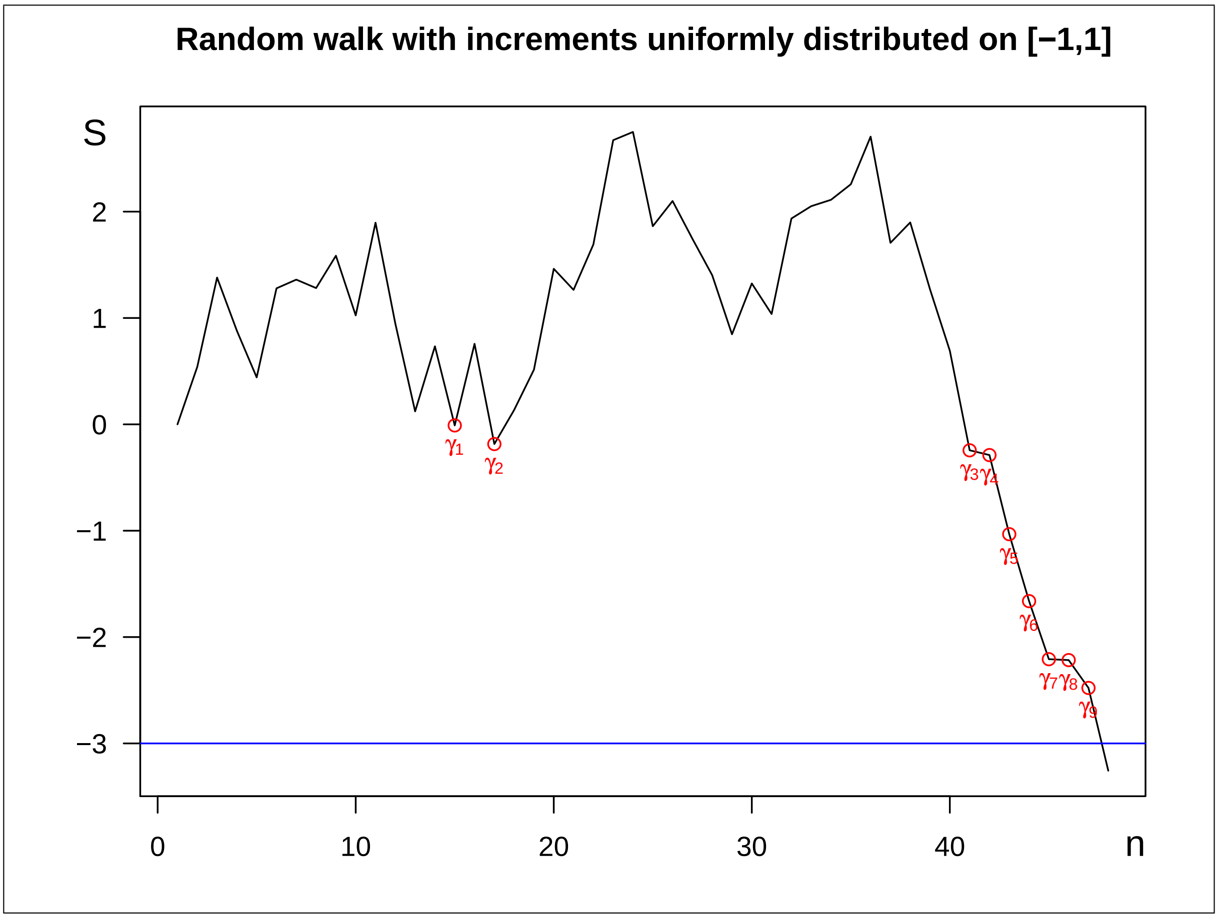


Figure 3.4: Example: Ladder points γ_i above level -3 for a random walk with increments that are uniformly distributed on $[-1, 1]$.

We refrain from indicating the dependence on n in our notation. Now by duality, S_n is a weakly ascending ladder height if

$$\{S_n \geq S_0, \dots, S_n \geq S_{n-1}\} \quad = \quad \{\hat{S}_n \geq 0, \dots, \hat{S}_1 \geq 0\} \;=\; \{\hat{L}_n \geq 0\} \;,$$

which proves the equivalence of (3.2) and (3.3).

For any oscillating random walk, both u and v are harmonic functions (see [BD94] and [AGKV05a]), that is

$$\begin{aligned} \mathbf{E}\big[u(x+X); X+x \geq 0\big] &= u(x)\;, \quad x \geq 0\;, \\ \mathbf{E}\big[v(x+X); X+x < 0\big] &= v(x)\;, \quad x \leq 0\;. \end{aligned} \tag{3.5}$$

Thus u and v can be used to construct new probability measures $\mathbf{P}^+$ and $\mathbf{P}^-$. The construction of these measures is described in detail in [AGKV05a]. For the sake of completeness, we recall the construction for $\mathbf{P}^-$. In the sequel, we denote by $\mathbf{P}_x$ and $\mathbf{E}_x$ that the random walk starts in $S_0 = x$.

For this, assume that S is adapted to some filtration $(\mathcal{F}_n)_{n\geq 0}$ such that X_{n+1} is independent of $\mathcal{F}_n$ for all $n \geq 1$. Let $R_0, R_1, \dots$ be a sequence of random variables with values in some state space $\mathcal{S}$, also adapted to $\mathcal{F}$. The sequence

$$v(S_0)\;,\; v(S_1)\mathbb{1}_{\{M_1<0\}}\;,\; v(S_2)\mathbb{1}_{\{M_2<0\}}\;,\;\dots$$

forms a martingale with respect to $\mathbf{P}_x$ with $x \leq 0$ and the filtration $\mathcal{F}$. Because of (3.5), for any bounded and measurable function $g : \mathcal{S}^{n+1} \to \mathbb{R}$,

$$\begin{aligned} &\mathbf{E}_x[g(R_0, \ldots, R_n)v(S_{n+1}); M_{n+1} < 0] \\ &\quad = \mathbf{E}_x[g(R_0, \ldots, R_n)v(S_n + X_{n+1}); M_n < 0, X_{n+1} < -S_n] \\ &\quad = \mathbf{E}_x[g(R_0, \ldots, R_n)v(S_n); M_n < 0] \ . \end{aligned}$$

This consistency property allows (under suitable regularity conditions on the underlying probability space) the construction of probability measures $\mathbf{P}_x^-$, $x \le 0$ fulfilling for each n

$$\mathbf{E}_x^-[g(R_0, \ldots, R_n)] \ = \ \frac{1}{v(x)} \mathbf{E}_x\big[g(R_0, \ldots, R_n)v(S_n); M_n < 0\big] \ .$$

The transformation above is known as the ***Doob transform*** from the theory of Markov chains. Under $\mathbf{P}_x^-$ the process $S_0, S_1, \ldots$ becomes a Markov chain with state space $\mathbb{R}^-$ and transition kernel

$$P^-(x, dy) \ := \ \frac{1}{v(x)} \mathbf{P}\{x + X \in dy\} v(y) 1_{\{y<0\}} \ , \quad x \le 0 \ .$$

As $P^-(x, [0, \infty)) = 0$, the Markov process described by the transition matrix above never enters $[0, \infty)$ again, although it may start from the boundary $x = 0$. Essentially, the transformation above describes a random walk conditioned to stay negative.

Similarly u gives rise to probability measures $\mathbf{P}_x^+$, $x \ge 0$, characterized by the equation

$$\mathbf{E}_x^+[g(R_0, \ldots, R_n)] \ = \ \frac{1}{u(x)} \mathbf{E}_x[g(R_0, \ldots, R_n)u(S_n); L_n \ge 0] \ , \ n \in \mathbb{N}_0 \ .$$

Under $\mathbf{P}^+$ $S_0, S_1, \ldots$ is a Markov process with state space $\mathbb{R}_0^+$ and transition probabilities

$$P^+(x, dy) \ := \ \frac{1}{u(x)} \mathbf{P}\{x + X \in dy\} u(y) 1_{\{y \ge 0\}} \ , \quad x \ge 0 \ .$$

Intuitively, it is the random walk conditioned to stay nonnegative.

Remark. *There is a slight difference between* $\mathbf{P}^+$ *and* $\mathbf{P}^-$*: under* $\mathbf{P}_x^+$ *the process* S *may hit 0, however, under* $\mathbf{P}_x^-$ *this possibility is excluded. For* $x < 0$ *a difference only occurs for those* x*, where* $v(x) \ne v(x-)$*, that is for at most countably many* x*. If one considers (as below) measures* $\mathbf{P}_\nu^-$ *having an initial distribution* ν *without atoms, there is no difference at all.*

3.2.2 A conditional limit law

Here, we prove a conditional limit law for oscillating random walks. It is a generalization of [BD94, Theorem 1]. In the context of BPREs the ideas of the proof of the limit theorem can be found in [Vat04, Lemma 7]. We use similar arguments here, but in a more general context. The limit theorem is valid under a more general condition than Assumption 3.2, often referred to as Spitzer's condition (see Chapter 2).

Assumption 3.5. *There exists a number* $0 < \rho < 1$ *such that*

$$\frac{1}{n} \sum_{k=1}^{n} \mathbf{P}(S_k > 0) \quad \to \quad \rho \quad \text{as } n \to \infty \ .$$

The summands may also be replaced by $\mathbf{P}(S_k \ge 0)$, as for every nondegenerated random walk, $\sum_{k=1}^n \mathbf{P}(S_k = 0) = o(n)$ (see [Fel87, chapter XII]). Note that Assumption 3.2 implies 3.5.

We will need the asymptotic of $\mathbf{P}(\tau_n = n)$. The next result follows from applying [AGKV05a, Lemma 2.1] to $-S_n$. The first part of the lemma can also be found in [BGT87].

Lemma 3.2.1. *Under Assumption 3.5, there are a slowly varying sequence $b_1, b_2, \dots$ and $\rho \le \alpha^{-1}$ such that for every $x < 0$*

$$\mathbf{P}_x(M_n < 0) \;=\; \mathbf{P}(M_n < -x) \;\sim\; v(x) b_n n^{-\rho}$$

and a constant $0 < d_3 < \infty$ such that

$$\mathbf{P}_x(M_n < 0) \;=\; \mathbf{P}(M_n < -x) \;\le\; d_3 v(x) b_n n^{-\rho} \;.$$

Note that by duality $\mathbf{P}(M_n < 0) = \mathbf{P}(\tau_n = n)$.

The slowly varying sequence (b_n) can be identified (see [BGT87, p. 382]):

$$b_n \;=\; \big(h(1-n^{-1}) \cdot \Gamma(1-\rho)\big)^{-1} \;,$$

where Γ is the Gamma-function and

$$h(s) \;=\; \exp\Big(\sum_{k=1}^{\infty} \frac{s^k}{k} \mathbf{P}(S_k \ge 0) - \rho\Big) \;.$$

The following proposition has been proved in [Vat04, Lemma 7] for a special function. A similar version for the measure $\mathbf{P}^+$ can be found in [AGKV05a, Lemma 2.5]. Here, we prove the corresponding version for $\mathbf{P}^-$. Recall from the preceding section that $R_0, R_1, \dots$ is a sequence of independent random variables adapted to the filtration $\mathcal{F}$.

Proposition 3.2.2. *Let $(\mathcal{U}_n)_{n\in\mathbb{N}}$ be a sequence of random variables in some Polish space S such that $\mathcal{U}_n$ is $\sigma(R_1, \dots, R_n)$ measurable for all n. Assume*

$$\mathcal{U}_n \;\to\; \mathcal{U}_\infty \quad \mathbf{P}^-\text{-a.s.} \tag{3.6}$$

for some S-valued random variable $\mathcal{U}_\infty$. Then under Assumption 3.5, for any continuous, bounded function $f : S \to \mathbb{R}$,

$$\lim_{n\to\infty} \mathbf{E}[f(\mathcal{U}_n) | M_n < 0] \;=\; \mathbf{E}^-[f(\mathcal{U}_\infty)] \;.$$

The dual version looks like the following:

Proposition 3.2.3. *Let $\mathcal{U}_n = \psi_n(R_1, \dots, R_n)$ and $\hat{\mathcal{U}}_n = \psi_n(R_n, \dots, R_1)$ be random variables with values in some Polish space S such that*

$$\hat{\mathcal{U}}_n \;\to\; \mathcal{U}_\infty \quad \mathbf{P}^-\text{-a.s.}$$

Then for any continuous, bounded function $f : S \to \mathbb{R}$

$$\lim_{n\to\infty} \mathbf{E}[f(\mathcal{U}_n) | \tau_n = n] \;=\; \mathbf{E}^-[f(\mathcal{U}_\infty)] \;.$$

The last proposition is an immediate consequence of

$$\mathbf{E}[f(\mathcal{U}_n) | \tau_n = n] \;=\; \frac{\mathbf{E}[f(\mathcal{U}_n); \tau_n = n]}{\mathbf{P}(\tau_n = n)} \;=\; \frac{\mathbf{E}[f(\hat{\mathcal{U}}_n); M_n < 0]}{\mathbf{P}(M_n < 0)} \;.$$

Proof of Proposition 3.2.2. We closely follow the ideas of [AGKV05a]. For the intermediately subcritical case, these ideas similar arguments can be found in [Vat04]. First, we show that

$$\lim_{n\to\infty} \mathbf{E}[f(\mathcal{U}_k) | M_n < 0] \;=\; \mathbf{E}^-[f(\mathcal{U}_k)] \;.$$

As the increments of the random walk are independent,

$$\begin{aligned}\mathbf{E}[f(\mathcal{U}_k)|M_n<0] &= \mathbf{P}(M_n<0)^{-1}\mathbf{E}[f(\mathcal{U}_k);\max\{S_{k+1}-S_k,\dots,S_n-S_k\}<-S_k,M_k<0] \\ &= \mathbf{P}(M_n<0)^{-1}\mathbf{E}[f(\mathcal{U}_k)\mathbf{P}_{S_k}(M_{n-k}<0);M_k<0]\ .\end{aligned}$$

By Lemma 3.2.1 and dominated convergence,

$$\lim_{n\to\infty}\mathbf{E}[f(\mathcal{U}_k)|M_n<0] \;=\; \mathbf{E}[f(\mathcal{U}_k)v(S_k);M_k<0] \;=\; \mathbf{E}^-[f(\mathcal{U}_k)]\ .$$

Note that for fixed k, the convergence also holds if f is bounded and measurable but not continuous.

Next, by the triangle inequality, for any $k\in\mathbb{N}$,

$$\begin{aligned}\Big|\mathbf{E}\big[f(\mathcal{U}_n)\big|M_n<0\big]-\mathbf{E}^-\big[f(\mathcal{U}_\infty)\big]\Big| &\le \mathbf{E}\big[|f(\mathcal{U}_n)-f(\mathcal{U}_k)|\big|M_n<0\big] \\ &\quad+\Big|\mathbf{E}\big[f(\mathcal{U}_k)\big|M_n<0\big]-\mathbf{E}^-\big[f(\mathcal{U}_\infty)\big]\Big|\ .\end{aligned}$$

As all functions here are bounded, the limits interchange with expectations. As we have proven above, the second term on the right-hand side converges to zero as n and then k go to infinity. It remains to prove that

$$\lim_{k\to\infty}\limsup_{n\to\infty}\mathbf{E}\Big[|f(\mathcal{U}_k)-f(\mathcal{U}_n)|\Big|M_n<0\Big] \;=\; 0\ .$$

For any $\delta>1$,

$$\begin{aligned}\mathbf{E}\Big[|f(\mathcal{U}_k)-f(\mathcal{U}_n)|\Big|M_n<0\Big] &= \mathbf{P}(M_n<0)^{-1}\mathbf{E}\Big[|f(\mathcal{U}_k)-f(\mathcal{U}_n)|;M_n<0\Big] \\ &\le \mathbf{P}(M_n<0)^{-1}\mathbf{E}\Big[|f(\mathcal{U}_k)-f(\mathcal{U}_n)|;M_{\lfloor\delta n\rfloor}<0\Big] \\ &\quad +2\ \sup\{f\}\ \mathbf{P}(M_n<0)^{-1}\mathbf{P}(M_n<0,M_{\lfloor\delta n\rfloor}\ge 0)\ .\end{aligned}\tag{3.7}$$

By Lemma 3.2.1

$$\begin{aligned}\mathbf{E}\Big[|f(\mathcal{U}_k)-f(\mathcal{U}_n)|;M_{\lfloor\delta n\rfloor}<0\Big] &= \mathbf{E}\Big[|f(\mathcal{U}_k)-f(\mathcal{U}_n)|\mathbf{P}(M_{\lfloor n(\delta-1)\rfloor}<-S_n);M_n<0\Big] \\ &\le \frac{d_3 b_{\lfloor n(\delta-1)\rfloor}}{\lfloor n(\delta-1)\rfloor^\rho}\mathbf{E}\Big[|f(\mathcal{U}_k)-f(\mathcal{U}_n)|v(-S_n);M_n<0\Big] \\ &= \frac{d_3 b_{\lfloor n(\delta-1)\rfloor}}{\lfloor n(\delta-1)\rfloor^\rho}\mathbf{E}^-\Big[|f(\mathcal{U}_k)-f(\mathcal{U}_n)|\Big]\ .\end{aligned}$$

Thus, again by Lemma 3.2.1 and properties of slowly varying functions (namely, $\lim_{n\to\infty}b_{\lfloor n\delta\rfloor}b_n^{-1}=1$; see appendix),

$$\limsup_{n\to\infty}\mathbf{P}(M_n<0)^{-1}\mathbf{E}\big[|f(\mathcal{U}_k)-f(\mathcal{U}_n)|;M_{\lfloor\delta n\rfloor}<0\big] \;\le\; d_3(\delta-1)^{-\rho}\mathbf{E}^-\big[|f(\mathcal{U}_k)-f(\mathcal{U}_\infty)|\big]\ .$$

By assumption, $f(\mathcal{U}_k)\to f(\mathcal{U}_\infty)$ $\mathbf{P}^-$-a.s. Thus, by dominated convergence, for any $\delta>1$, the above term converges to zero as $k\to\infty$. The other term in (3.7) can be treated as follows:

$$\begin{aligned}\lim_{n\to\infty}\mathbf{P}(M_n<0)^{-1}\mathbf{P}(M_n<0,M_{\lfloor\delta n\rfloor}\ge 0) &= 1-\lim_{n\to\infty}\frac{\mathbf{P}(M_{\lfloor\delta n\rfloor}<0)}{\mathbf{P}(M_n<0)} \\ &= 1-\delta^{-\rho}\end{aligned}$$

Taking the limit $\delta\to 1$ yields the desired result. □

A large deviation result we will use later can be found in [ABKV10, Proposition 1 and Corollary 2.4].

Lemma 3.2.4. *Under Assumption 3.2, for all $r>0$,*

$$\mathbf{E}_x[e^{-rS_n};L_n\ge 0] \;\sim\; s(0)\ n^{-1}\ a_n^{-1}\ u(x)\ \int_0^\infty e^{-rz}v(-z)dz \tag{3.8}$$

and there is a $c>0$ such that

$$\mathbf{E}_x[e^{-rS_n};L_n\ge 0] \;\le\; c\ n^{-1}\ a_n^{-1}\ u(x)\ . \tag{3.9}$$

3.3 Auxiliary results for the BPRE

Let

$$\eta_k \quad := \quad \sum_{y=0}^{\infty} y(y-1)Q_k(\{y\}) \Big/ m(Q_k)^2, \quad k \geq 1$$

be the standardized second factorial moment of the offspring distribution in generation k.

Lemma 3.3.1. *Under Assumption 3.2 and 3.3,*

$$\sum_{k=1}^{\infty} \eta_k e^{S_k} \quad < \quad \infty \quad \mathbf{P}^- \text{-}a.s.$$

Additionally, for all continuity points $c > 0$ of the distribution of $\sum_{k=1}^{\infty} \eta_k e^{S_k}$,

$$\lim_{n\to\infty} \mathbf{P}\left(\sum_{k=1}^{n} \eta_k e^{S_k} < c \Big| M_n < 0 \right) \quad = \quad \mathbf{P}^- \left(\sum_{k=1}^{\infty} \eta_k e^{S_k} < c \right) \, . \tag{3.10}$$

The first statement is proved just the same way as a similar statement for the measure $\mathbf{P}^+$ that can be found in [AGKV05a, Lemma 2.7] and [ABKV10, Lemma 3.1]. The second statement is an application of Proposition 3.2.2, as $\mathbf{1}_{\{\sum_{k=1}^{n} \eta_k e^{S_k} < c\}}$ is a bounded and $\sigma(Q_1, \dots, Q_n)$-measurable function. For any continuity point of the distribution of $\sum_{k=1}^{\infty} \eta_k e^{S_k}$, it is also $\mathbf{P}^-$-a.s. continuous. Note that there are at most countably many points of discontinuity and therefore we can always find an abritrarily large c such that (3.10) is fulfilled.

Let

$$\begin{aligned} U_n &:= e^{-S_n} \\ V_n &:= U_n + \sum_{k=0}^{n-1} \eta_{k+1} U_k \, . \end{aligned}$$

It turns out that U_n and V_n already describe some key properties of Z_n, conditioned on Π. In particular, they provide a lower bound for the survival probability. We use a formula derived in [GK00]. Let

$$f_n(s) \quad := \quad \sum_{k=0}^{\infty} s^k Q_n(\{k\})$$

be the probability generating function of the offspring distribution of an individual in generation $n-1$. Note that $X_n = \log f_n'(1)$. Then, as it is well-known (see e.g. [AN72]),

$$\mathbb{E}[s^{Z_n}|\Pi] \quad = \quad f_1(f_2(\cdots f_n(s)\cdots)) \; = \; f_{0,n}(s), \quad s \geq 0. \tag{3.11}$$

We will use an alternative expression for this generating function. Let

$$\begin{aligned} f_{k,n} &:= f_{k+1} \circ f_{k+2} \circ \cdots \circ f_n, \quad 0 \leq k < n; \; f_{n,n} = id \\ g_k(s) &:= \frac{1}{1-f_k(s)} - \frac{1}{f_k'(1)(1-s)}, \quad s \geq 0 \, , \end{aligned} \tag{3.12}$$

where id denotes the indentity function.

As $X_n = \log f_n'(1)$, U_k may be written as

$$U_k \quad = \quad (f_1'(1) \cdots f_k'(1))^{-1} \; = \; f_{0,k}'(1)^{-1} \, .$$

By a telescope summation argument,

$$\begin{aligned}
\frac{1}{1-f_{0,n}(s)} &= \frac{U_0}{1-f_{0,n}(s)} \\
&= \frac{U_n}{1-f_{n,n}(s)} + \sum_{k=0}^{n-1}\left(\frac{U_k}{1-f_{k,n}(s)} - \frac{U_{k+1}}{1-f_{k+1,n}(s)}\right) \\
&= \frac{U_n}{1-s} + \sum_{k=0}^{n-1} U_k \left(\frac{1}{1-f_{k+1}(f_{k+1,n}(s))} - \frac{1}{f'_{k+1}(1)(1-f_{k+1,n}(s))}\right) \\
&= \frac{U_n}{1-s} + \sum_{k=0}^{n-1} U_k g_{k+1}(f_{k+1,n}(s)), \quad s \geq 0. \qquad (3.13)
\end{aligned}$$

Note that (3.13) does not only hold for those s in the domain of convergence of $f_{0,n}(s)$, but for all $s \geq 0$. In [GK00] it is proved that for $s \in [0,1)$

$$g(s) \;\leq\; \frac{f''(1)}{\big(f'(1)\big)^2} \;=\; \eta\;.$$

Thus taking $s=0$ in (3.13) yields the following estimate, already found in [Agr75]

$$\mathbb{P}(Z_n > 0|\Pi) \;=\; 1 - f_{0,n}(0) \;\geq\; V_n^{-1}\;.$$

Together with (1.3),

$$V_n^{-1} \;\leq\; \mathbb{P}(Z_n > 0|\Pi) \;\leq\; \exp(L_n) \quad \text{a.s.} \qquad (3.14)$$

The following lemma says that U_n and V_n are of the same order if S_n has its minimum at the end.

Lemma 3.3.2. *Under Assumptions 3.1 to 3.3, there is a $d_1 > 0$ such that for all $n \in \mathbb{N}$,*

$$\mathbf{E}\left[\frac{U_n}{V_n}\Big|\tau_n = n\right] \;\geq\; d_1\;.$$

Proof. It suffices to show

$$\mathbf{P}(U_n/V_n > c, \tau_n = n) \;\geq\; d\,\mathbf{P}(\tau_n = n) \qquad (3.15)$$

for some $d, c > 0$ small enough. By duality,

$$\begin{aligned}
\mathbf{P}(U_n/V_n > c, \tau_n = n) &= \mathbf{P}\left(\sum_{k=0}^{n-1}\eta_{k+1}e^{S_n - S_k} < c^{-1}-1, \tau_n = n\right) \\
&= \mathbf{P}(M_n < 0)\,\mathbf{P}\left(\sum_{k=1}^{n}\eta_k e^{S_k} < c^{-1}-1\Big|M_n < 0\right)\;.
\end{aligned}$$

By Lemma 3.3.1, $\mathbf{P}\left(\sum_{k=1}^{n}\eta_k e^{S_k} < c^{-1}-1\Big|M_n < 0\right)$ converges to $\mathbf{P}^-(\sum_{k=1}^{\infty}\eta_k e^{S_k} < c^{-1}-1)$ for suitable c. Also by Lemma 3.3.1

$$\sum_{k=1}^{\infty}\eta_k e^{S_k} \;<\; \infty \qquad \mathbf{P}^-\text{–a.s.}$$

Thus for $c > 0$ and $d > 0$ small enough, $\mathbf{P}^-(\sum_{k=1}^{\infty}\eta_k e^{S_k} < c^{-1}-1) > 2d$. Now choosing n^* large enough,

$$\mathbf{P}\left(\sum_{k=1}^{n}\eta_k e^{S_k} < c^{-1}-1\Big|M_n < 0\right) \geq d$$

for all $n > n^*$. Thus we have the desired estimate for $n > n^*$.

For $n \le n^*$,

$$\min_{0\le n\le n^*} \mathbb{E}\Big[\frac{U_n}{V_n}\Big|\tau_n = n\Big] \ > \ 0$$

and thus the constant defined by

$$d_1 \ := \ \min\Big\{d, \min_{0\le n\le n^*} \mathbb{E}\Big[\frac{U_n}{V_n}\Big|\tau_n = n\Big]\Big\}$$

fulfills the requirements of the lemma. $\square$

The next Lemma provides a precise mathematical statement for the heuristics described on p. 5:

Lemma 3.3.3. *Under Assumptions 3.1 and 3.2, there is a $d_2 < \infty$ such that uniformly in n*

$$\mathbf{E}\left[e^{L_n - S_n}\right] \ \le \ d_2 \mathbf{P}(\tau_n = n) \ .$$

Proof. We make a decomposition according to the first minimum of the random walk. By Markov property and monotonicity of $\mathbf{P}(\tau_n = n) = \mathbf{P}(M_n < 0)$,

$$\begin{aligned}
\mathbf{E}\left[e^{L_n - S_n}\right] \ &= \ \sum_{k=0}^{n} \mathbf{E}\left[e^{-(S_n - S_k)}; \tau_n = k\right] \\
&= \ \sum_{k=0}^{n} \mathbf{E}\left[e^{-(S_n - S_k)}; \tau_k = k, \min(S_k, \ldots, S_n) \ge S_k\right] \\
&= \ \sum_{k=0}^{n} \mathbf{P}(\tau_k = k) \ \mathbf{E}\left[e^{-S_{n-k}}; L_{n-k} \ge 0\right] \\
&\le \ \sum_{k=0}^{\lfloor n/2\rfloor} \mathbf{P}(\tau_k = k)\mathbf{E}\left[e^{-S_{n-k}}; L_{n-k} \ge 0\right] \\
&\quad + \mathbf{P}\left(\tau_{\lceil n/2\rceil} = \lceil n/2\rceil\right) \sum_{k=0}^{\lceil n/2\rceil} \mathbf{E}\left[e^{-S_k}; L_k \ge 0\right] \ .
\end{aligned}$$

By Lemma 3.2.4,

$$\sum_{k=0}^{\infty} \mathbf{E}\left[e^{-S_k}; L_k \ge 0\right] \ < \ \infty$$

and thus, again by Lemma 3.2.4, for $c < \infty$ large enough

$$\mathbf{E}\left[e^{L_n - S_n}\right] \ \le \ c\lfloor n/2\rfloor^{-1} a_{\lfloor n/2\rfloor}^{-1} \sum_{k=0}^{\lfloor n/2\rfloor} \mathbf{P}(\tau_k = k) + c\mathbf{P}\left(\tau_{\lceil n/2\rceil} = \lceil n/2\rceil\right)$$

By Lemma 3.2.1 $\mathbf{P}(\tau_n = n) = \mathbf{P}(M_n < 0)$ is regularly varying with exponent $-\rho \in (-1, 0)$. An application of Karamata's theorem (see appendix, Theorem A.2.1) yields

$$\sum_{k=0}^{\lfloor n/2\rfloor} \mathbf{P}(\tau_k = k) \ \sim \ \frac{n/2}{1-\rho}\mathbf{P}(\tau_{\lfloor n/2\rfloor} = \lfloor n/2\rfloor)$$

and, together with Lemma 3.2.1,

$$n^{-1}\mathbf{P}(\tau_n = n)^{-1} \sum_{k=0}^{\lfloor n/2\rfloor} \mathbf{P}(\tau_k = k) \ \to \ \frac{1}{2^{1-\rho}(1-\rho)} \ .$$

Thus for every $\epsilon > 0$ and some n^* large enough, for all $n \ge n^*$

$$n^{-1}\mathbf{P}(\tau_n = n)^{-1} \sum_{k=0}^{\lfloor n/2\rfloor} \mathbf{P}(\tau_k = k) \ \le \ (1+\epsilon)\frac{1}{2^{1-\rho}(1-\rho)}$$

and

$$\frac{\mathbf{P}(\tau_{\lfloor n/2 \rfloor} = \lfloor n/2 \rfloor)}{\mathbf{P}(\tau_n = n)} \quad \le \quad (1+\epsilon)1/2^\rho \ .$$

Therefore, for all $n \ge n^*$, we can estimate, with $c < \infty$ large enough,

$$\mathbf{E}\left[e^{L_n - S_n}\right] \quad \le \quad c\Big(2a_{\lfloor n/2 \rfloor}^{-1}(1+\epsilon)\frac{1}{2^{1-\rho}(1-\rho)} + 1\Big)\mathbf{P}(\tau_n = n) \ .$$

As a_n^{-1} converges to zero as $n \to \infty$, this yields the result for $n \ge n^*$.

For $n \le n^*$, we have $\mathbf{P}(\tau_n = n) = \mathbf{P}(M_n < 0) \ge \mathbf{P}(X < 0)^n \ge \mathbf{P}(X < 0)^{n^*}$. The result follows by taking d_2 larger than $\mathbf{P}(X < 0)^{-n^*}$ (since $\mathbf{E}[e^{L_n - S_n}] \le 1$). □

3.4 Proof of Theorems

By the previous lemmata, we immediately obtain bounds uniformly in n:

Theorem 3.4.1. *Under Assumptions 3.1 to 3.3, there are constants $c_1, c_2 > 0$ such that uniformly in n*

$$\mathbb{P}(Z_n > 0) \quad \le \quad c_1 \gamma^n \mathbf{P}(\tau_n = n)$$

and

$$\mathbb{P}(Z_n > 0) \quad \ge \quad c_2 \gamma^n \mathbf{P}(\tau_n = n) \ .$$

Proof of Theorem 3.4.1. By (3.14)

$$\mathbb{P}(Z_n > 0) \quad \ge \quad \mathbb{E}[V_n^{-1}] \ = \ \gamma^n \mathbf{E}\left[\frac{U_n}{V_n}\Big|\tau_n = n\right]\mathbf{P}(\tau_n = n)$$

and

$$\mathbb{P}(Z_n > 0) \quad \le \quad \gamma^n \mathbf{E}\left[e^{L_n - S_n}\right] \ .$$

Together with Lemmata 3.3.2 and 3.3.3, this yields the theorem. □

3.4.1 Proof of Theorem 3.1.1

For the proof, we make a change of measure and apply the conditional limit law for the recurrent walk:

Lemma 3.4.2. *Under Assumptions 3.1 to 3.3,*

$$\mathbf{E}\big[e^{-S_n}\mathbf{P}(Z_n > 0|\Pi)|\tau_n = n\big] \quad \to \quad \mathbf{E}^-[U_\infty]$$

where the random variable U_∞ is $\mathbf{P}^-$–a.s. strictly positive.

Proof. Define similarly to (3.11)

$$f_{n,0}(s) \quad := \quad f_n \circ \cdots \circ f_1(s) \ . \tag{3.16}$$

Let

$$\mathcal{U}_n(s) \ := \ e^{-S_n}(1 - f_{0,n}(s)) \ . \tag{3.17}$$

Then $\mathcal{U}_n(0) = e^{-S_n}\mathbf{P}(Z_n > 0|\Pi)$. The corresponding dual process is $\hat{\mathcal{U}}_n(s) = e^{-S_n}(1 - f_{n,0}(s))$. We show $\mathbf{P}^-$–a.s. convergence of $\hat{\mathcal{U}}_n$ by similar arguments as in [Vat04]. By convexity, for $s \in [0,1]$,

$$f_{n+1,0}(s) \ = \ f_{n+1}(f_{n,0}(s)) \quad \ge \quad 1 - (1 - f_{n,0}(s))f'_{n+1}(1) \ .$$

Thus

$$\begin{aligned}
\hat{\mathcal{U}}_{n+1}(s) &= e^{-S_n} e^{-X_{n+1}} (1 - f_{n+1}(f_{n,0}(s))) \\
&\leq e^{-S_n} (1 - f_{n,0}(s)) e^{-X_{n+1}} f'_{n+1}(1) \ = \ \hat{\mathcal{U}}_n(s) \quad \mathbf{P}^-\text{–a.s.}
\end{aligned}$$

and $(\hat{\mathcal{U}}_n)_{n\geq 1}$ is a nonnegative, nonincreasing sequence in n. Thus, $\hat{\mathcal{U}}_n$ converges $\mathbf{P}^-$–a.s. to a random variable $\mathcal{U}_\infty$ bounded by 1. The convergence result then follows from Proposition 3.2.3. $\mathcal{U}_\infty$ being strictly positive follows rom (3.14) and Lemma 3.3.1. □

Proof of Theorem 3.1.1. For the proof, we make a decomposition according to the minimum of the associated random walk. Let $m \in \mathbb{N}$. Then

$$\begin{aligned}
\mathbb{P}(Z_n > 0) &= \sum_{k=0}^{n} \mathbb{E}[\mathbb{P}(Z_n > 0 | \Pi); \tau_k = k, \min\{S_k, \ldots, S_n\} \geq S_k] \\
&= \sum_{k=0}^{n} \mathbb{E}[1 - f_{0,n}(0); \tau_k = k, \min\{S_k, \ldots, S_n\} \geq S_k] \\
&= \sum_{k=0}^{n} \gamma^n \mathbf{E}[e^{-S_n}(1 - f_{0,n}(0)); \tau_k = k, \min\{S_k, \ldots, S_n\} \geq S_k] \\
&= \sum_{k=0}^{n-m-1} \gamma^n \mathbf{E}[e^{-S_n}(1 - f_{0,n}(0)); \tau_k = k, \min\{S_k, \ldots, S_n\} \geq S_k] \\
&\qquad + \sum_{k=n-m}^{n} \gamma^n \mathbf{E}[e^{-S_n}(1 - f_{0,n}(0)); \tau_k = k, \min\{S_k, \ldots, S_n\} \geq S_k] \\
&=: \chi_1 + \chi_2 \quad .
\end{aligned}$$

The first sum can be estimated by

$$\begin{aligned}
\chi_1 &\leq \gamma^n \sum_{k=0}^{n-m-1} \mathbf{E}[e^{-S_{n-k}}; L_{n-k} \geq 0] \mathbf{P}(\tau_k = k) \\
&= \gamma^n \sum_{k=m+1}^{n} \mathbf{E}[e^{-S_k}; L_k \geq 0] \mathbf{P}(\tau_{n-k} = n-k) \ .
\end{aligned}$$

By Lemma 3.2.4 and monotonicity of $\mathbf{P}(\tau_n = n)$ (recall $a_n = n^{1/\alpha} l_n$),

$$\begin{aligned}
\chi_1 &\leq c\,\gamma^n \sum_{k=m+1}^{n} \frac{1}{a_k k} \mathbf{P}(\tau_{n-k} = n-k) \\
&\leq c\,\gamma^n \mathbf{P}(\tau_{\lfloor n/2 \rfloor} = \lfloor n/2 \rfloor) \sum_{k=m+1}^{\lfloor n/2 \rfloor} \frac{1}{a_k k} + c\,\gamma^n a_{\lfloor n/2 \rfloor}^{-1} \lfloor n/2 \rfloor^{-1} \sum_{k=0}^{\lfloor n/2 \rfloor} \mathbf{P}(\tau_k = k) \ .
\end{aligned}$$

In view of Theorem A.2.1, $\sum_{k=0}^{\lfloor n/2 \rfloor} \mathbf{P}(\tau_k = k) \sim c\, n \mathbf{P}(\tau_n = n)$ and thus, for some $c < \infty$,

$$\limsup_{n\to\infty} \gamma^{-n} \mathbf{P}(\tau_n = n)^{-1} \chi_1 \quad \leq \quad c \sum_{k=m}^{\infty} k^{-1} a_k^{-1} \ < \ \infty \ .$$

For the second sum, we condition on $\sigma(Q_k, \ldots, Q_n)$. By $^\wedge$, we denote independent copies of the corresponding random variable that are measurable with respect to the filtration $\hat{\Pi}_j := \sigma(\hat{Q}_1, \ldots, \hat{Q}_j)$, namely

$\hat{L}_j$, $\hat{S}_j$, and $\hat{A}_j := \hat{f}_{0,j}(0)$. This yields

$$\begin{aligned}\chi_2 = & \sum_{k=n-m}^{n} \gamma^n \mathbf{E}\big[\mathbf{E}[e^{-S_k}(1-f_{0,k}(f_{k+1,n}(0)))|\tau_k = k, Q_k, \ldots, Q_n]e^{-(S_n-S_k)} \\ & \qquad\qquad ;\min\{S_{k+1}-S_k, \ldots, S_n - S_k\} \geq 0\big] \cdot \mathbf{P}(\tau_k = k) \\ = & \sum_{j=0}^{m} \gamma^n \mathbf{E}\big[\mathbf{E}[e^{-S_{n-j}}(1-f_{0,n-j}(\hat{A}_j))|\tau_{n-j} = n-j, \hat{\Pi}_j]e^{-\hat{S}_j}; \hat{L}_j \geq 0\big] \cdot \mathbf{P}(\tau_{n-j} = n-j) \ .\end{aligned}$$

We set

$$s_{n-j}(\hat{A}_j) \quad := \quad \mathbf{E}\big[e^{-S_{n-j}}(1-f_{0,n-j}(\hat{A}_j))|\tau_{n-j} = n-j, \hat{\Pi}_j\big] \ .$$

In the proof of Lemma 3.4.2. it has been shown that the dual process of $e^{-S_k}(1-f_{0,k}(s))$ is nonincreasing and converges $\mathbf{P}^-$–a.s. By $\mathcal{U}_\infty(s)$ we denote its limit. Thus, by Proposition 3.2.3, for fixed j, $s_{n-j}(\hat{A}_j)$ converges to $\mathcal{U}_\infty^-(\hat{A}_j) := \mathbf{E}^-[\mathcal{U}_\infty(\hat{A}_j)|\hat{\Pi}_j]$. As $(\hat{Q}_1, \ldots, \hat{Q}_j)$ are identical in distribution to $(Q_1, \ldots, Q_j)$, the $^\wedge$ may be dropped in the following. Thus, again by Lemma 3.2.1,

$$\lim_{n\to\infty} \gamma^{-n}\mathbf{P}(\tau_n = n)^{-1}\chi_2 \quad = \quad \sum_{j=0}^{m} \mathbf{E}\big[\mathcal{U}_\infty^-(A_j)e^{-S_j}; L_j \geq 0\big] \ .$$

Taking the limit $m \to \infty$, $\limsup_{n\to\infty} \gamma^{-n}\mathbf{P}(\tau_n = n)^{-1}\chi_1$ vanishes and

$$\lim_{n\to\infty} \frac{\mathbb{P}(Z_n > 0)}{\gamma^n \mathbf{P}(\tau_n = n)} \quad = \quad \sum_{j=0}^{\infty} \mathbf{E}\big[\mathcal{U}_\infty^-(A_j)e^{-S_j}; L_j \geq 0\big] \quad =: \ \theta \ , \tag{3.18}$$

which proves the theorem. θ being strictly positive follows from Theorem 3.4.1. □

3.4.2 Proof of Theorem 3.1.2.

It is well-known (see e.g. [Afa93]) that

$$\mathbb{E}[Z_n(Z_n - 1)|\Pi] \quad \leq \quad \frac{V_n}{U_n^2} \quad \text{a.s.} \tag{3.19}$$

As we see from the last estimate, Z_n has bounded variance at times when U_n/V_n is large (as $1/V_n \leq 1$). In this case, the population won't be too large and there is a large probability that the process dies out in the next step. We start by showing that – conditioned on survival – the expected number of those times is of order $n^{1-\rho}$.

Lemma 3.4.3. *Under Assumptions 3.1 to 3.3, for $c > 0$ small enough there are constants $c_1, c_2 > 0$ such that*

$$\limsup_{n\to\infty} \frac{1}{n\mathbf{P}(\tau_n = n)} \sum_{k=0}^{n} \mathbb{E}\left[\mathbb{1}_{\{U_k/V_k > c\}}|Z_n > 0\right] \quad \leq \quad c_1$$

and

$$\liminf_{n\to\infty} \frac{1}{n\mathbf{P}(\tau_n = n)} \sum_{k=0}^{n} \mathbb{E}\left[\mathbb{1}_{\{U_k/V_k > c\}}|Z_n > 0\right] \quad \geq \quad c_2 \ .$$

Proof of lemma. First we prove the lower bound. We use (3.14) and change to the measure $\mathbf{P}$ to make use of properties of oscillating random walks,

$$\begin{aligned}\sum_{k=0}^{n} \mathbb{E}[\mathbb{1}_{\{U_k/V_k > c\}}|Z_n > 0] \quad = & \quad \mathbb{P}(Z_n > 0)^{-1} \sum_{k=0}^{n} \mathbb{E}\left[\mathbf{P}(Z_n > 0|\Pi)\mathbb{1}_{\{U_k/V_k > c\}}\right] \\ \geq & \quad \mathbb{P}(Z_n > 0)^{-1}\gamma^n \sum_{k=0}^{n} \mathbf{E}[U_n/V_n; U_k/V_k > c] \ .\end{aligned}$$

We denote

$$\begin{aligned}
\hat{S}_j &:= S_{k+j} - S_k \ , \quad \hat{\eta}_i := \eta_{i+k} \\
\hat{U}_k &:= e^{-\hat{S}_k} \\
\hat{V}_k &:= \hat{U}_k + \sum_{i=0}^{k-1} \hat{\eta}_{i+1} e^{-\hat{S}_i}
\end{aligned}$$

and refrain from indicating the dependence on k in our notation. All random variables marked with $\wedge$ are independent of $\sigma(S_0, \ldots, S_k)$. Now by definition of V_n,

$$\begin{aligned}
\frac{U_n}{V_n} &= \frac{e^{-S_n}}{e^{-S_n} + \sum_{i=0}^{n-1} \eta_{i+1} e^{-S_i}} \\
&= \frac{e^{-S_n}}{e^{-S_n} + \sum_{i=0}^{k-1} \eta_{i+1} e^{-S_i} + e^{-S_k} \sum_{i=k}^{n-1} \eta_{i+1} e^{S_k - S_i}} \\
&= \frac{e^{-S_n}}{\sum_{i=0}^{k-1} \eta_{i+1} e^{-S_i} + e^{-S_k} (e^{-(S_n - S_k)} + \sum_{i=k}^{n-1} \eta_{i+1} e^{S_k - S_i})} \\
&= \frac{e^{-(S_n - S_k)}}{e^{S_k} \sum_{i=0}^{k-1} \eta_{i+1} e^{-S_i} + (e^{-(S_n - S_k)} + \sum_{i=k}^{n-1} \eta_{i+1} e^{S_k - S_i})} \\
&= \frac{\hat{U}_{n-k}}{V_k/U_k - 1 + \hat{V}_{n-k}} \ .
\end{aligned}$$

Therefore,

$$\begin{aligned}
\mathbf{E}[U_n/V_n; U_k/V_k > c] &= \mathbf{E}\left[\frac{\hat{U}_{n-k}}{V_k/U_k - 1 + \hat{V}_{n-k}}; U_k/V_k > c\right] \\
&\geq \mathbf{P}(U_k/V_k > c)\mathbf{E}\left[\frac{U_{n-k}}{c^{-1} - 1 + V_{n-k}}\right] \ .
\end{aligned}$$

As all random variables with $\wedge$ are i.i.d. to the corresponding random variables, we dropped the $\wedge$ in the last line. Now, conditioned on $\{\tau_{n-k} = n-k\}$, $U_{n-k} \geq 1$ [15] and then

$$\frac{U_{n-k}}{c^{-1} - 1 + V_{n-k}} \geq \frac{c}{U_{n-k}^{-1} + cV_{n-k}/U_{n-k}} \geq \frac{c}{1 + cV_{n-k}/U_{n-k}}.$$

Thus, in view of (3.15), for $c, d > 0$ small enough

$$\begin{aligned}
\mathbf{E}[U_n/V_n; U_k/V_k > c] &\geq \frac{c}{2}\mathbf{P}(U_k/V_k > c)\mathbf{P}(U_{n-k}/V_{n-k} > c, \tau_{n-k} = n-k) \\
&\geq \frac{c}{2}\mathbf{P}(U_k/V_k > c, \tau_k = k)\mathbf{P}(U_{n-k}/V_{n-k} > c, \tau_{n-k} = n-k) \\
&\geq d^2 \frac{c}{2}\mathbf{P}(\tau_k = k)\mathbf{P}(\tau_{n-k} = n-k) \ .
\end{aligned}$$

So the proof is reduced to the analysis of $\mathbf{P}(\tau_k = k)$. As $\mathbf{P}(\tau_k = k)$ is monotonly decreasing in k,

$$\sum_{k=0}^{n} \mathbf{P}(\tau_k = k)\mathbf{P}(\tau_{n-k} = n-k) \quad \geq \quad 2\mathbf{P}(\tau_n = n) \sum_{k=0}^{\lfloor n/2 \rfloor} \mathbf{P}(\tau_k = k) \ .$$

By Lemma 3.2.1, $\mathbf{P}(\tau_k = k) = k^{-\rho} b_k (1 + o(1))$. Thus, using Theorem A.2.1 from the appendix,

$$\sum_{k=0}^{\lfloor n/2 \rfloor} \mathbf{P}(\tau_k = k) \sim \frac{n}{2(1-\rho)}\mathbf{P}(\tau_{\lfloor n/2 \rfloor} = \lfloor n/2 \rfloor) \sim \frac{1}{2^{1-\rho}(1-\rho)} n\mathbf{P}(\tau_n = n) \ . \tag{3.20}$$

[15] since $S_0 = 0$ and therefore, $\{\tau_{n-k} = n-k\}$ implies $\{S_{n-k} \leq 0\}$

Therefore, by Theorem 3.4.1 and (3.20),

$$\liminf_{n\to\infty} \frac{\gamma^n \sum_{k=0}^n \mathbf{P}(\tau_k = k)\mathbf{P}(\tau_{n-k} = n-k)}{n\mathbf{P}(\tau_n = n)\mathbb{P}(Z_n > 0)} \;\geq\; c_2 \liminf_{n\to\infty} \frac{\mathbf{P}(\tau_n = n)\sum_{k=0}^{\lfloor n/2\rfloor} \mathbf{P}(\tau_k = k)}{n\mathbf{P}(\tau_n = n)\mathbf{P}(\tau_n = n)} \;>\; 0 \;,$$

which is the desired lower bound.

For the upper bound, we use (3.14) which also implies $V_n^{-1} \leq e^{L_n}$ a.s. Now $L_n \leq \min\{S_k, \ldots, S_n\}$ and thus

$$\begin{aligned}
\sum_{k=0}^n \mathbb{E}\left[\mathbb{1}_{\{U_k/V_k > c\}} | Z_n > 0\right] &\leq \mathbb{P}(Z_n > 0)^{-1}\gamma^n \sum_{k=0}^n \mathbf{E}\left[e^{L_n - S_n}; U_k/V_k > c\right] \\
&\leq \mathbb{P}(Z_n > 0)^{-1}\gamma^n \sum_{k=0}^n \mathbf{E}\Big[e^{\min\{S_k - S_k, \ldots, S_n - S_k\} - (S_n - S_k)}; U_k/V_k > c\Big] \\
&= \mathbb{P}(Z_n > 0)^{-1}\gamma^n \sum_{k=0}^n \mathbf{E}\left[e^{L_{n-k} - S_{n-k}}\right] \mathbf{P}(U_k/V_k > c) \;. \qquad (3.21)
\end{aligned}$$

By Markov-inequality, (3.14), and Lemma 3.3.3,

$$\mathbf{P}(U_{n-k}/V_{n-k} > c) \;\leq\; c^{-1}\mathbf{E}\Big[U_{n-k}/V_{n-k}\Big] \;\leq\; c^{-1}\mathbf{E}\Big[e^{L_{n-k} - S_{n-k}}\Big] \;\leq\; c^{-1}\, d_2\, \mathbf{P}(\tau_{n-k} = n-k) \;.$$

Thus,

$$\limsup_{n\to\infty} \frac{\sum_{k=0}^n \mathbb{E}\left[\mathbb{1}_{\{U_k/V_k > c\}} | Z_n > 0\right]}{n\mathbf{P}(\tau_n = n)} \;\leq\; \limsup_{n\to\infty} \frac{d_2}{c}\, \frac{\sum_{k=0}^n \mathbf{P}(\tau_k = k)\mathbf{P}(\tau_{n-k} = n-k)}{\gamma^{-n}\mathbb{P}(Z_n > 0)n\mathbf{P}(\tau_n = n)} \;.$$

Theorem 3.4.1 and an application of Theorem A.2.1 from the appendix yield the result. □

Proof of the upper bound in Theorem 3.1.2. Since

$$\begin{aligned}
\sum_{k=1}^n \mathbb{P}(Z_n > 0)^{-1}\mathbb{E}[\mathbb{1}_{\{Z_k = 1, Z_n > 0\}}] &\leq \sum_{k=1}^n \mathbb{P}(Z_n > 0)^{-1}\mathbb{E}[\mathbb{P}(Z_k > 0|\Pi)]\mathbb{E}[\mathbb{P}(Z_{n-k} > 0|\Pi)] \\
&\leq \mathbb{P}(Z_n > 0)^{-1}\gamma^n \sum_{k=1}^n \mathbf{E}\left[e^{L_k - S_k}\right]\mathbf{E}\left[e^{L_{n-k} - S_{n-k}}\right] \;,
\end{aligned}$$

the first part of the theorem follows by Lemma 3.3.3 and the same application of Karamata's theorem as in the previous proof. □

Proof of the lower bound in Theorem 3.1.2. In this case, it remains to prove that $U_n/V_n > c$ implies that, conditioned on survival, the probability of having just one individual in the next generation is bounded from below by some constant:

Lemma 3.4.4. *Under Assumptions 3.1 to 3.4, for any* $0 < c < 1$*, there exists a* $p > 0$ *such that for all* $k < n \in \mathbb{N}$

$$\mathbb{P}(Z_{k+1} = 1 | Z_n > 0, U_k/V_k > c) \;\geq\; p \;.$$

Thus conditioned on $\{U_n/V_n > c\}$, there is a positive probability that there is just one individual left in the next generation. Conditioned on survival in generation n, the number of those times is of order $\Theta(n^{1-\rho})$. We can estimate

$$\begin{aligned}
\sum_{k=1}^n \mathbb{E}\big[\mathbb{1}_{\{Z_k = 1, Z_n > 0\}}\big] &\geq \sum_{k=0}^{n-1} \mathbb{P}(Z_{k+1} = 1 | Z_n > 0, U_k/V_k > c)\mathbb{P}(U_k/V_k > c, Z_n > 0) \\
&\geq p \sum_{k=1}^n \mathbb{E}\big[\mathbb{1}_{\{U_{k-1}/V_{k-1} > c, Z_n > 0\}}\big] \;.
\end{aligned}$$

Together with Lemma 3.4.3 and Theorem 3.1.1, this proves the lower bound. □

Proof of Lemma 3.4.4. First note that in view of (3.14) for every $c > 0$,

$$\begin{aligned} \mathbb{P}(Z_n > 0, U_n/V_n > c) &= \mathbb{E}\big[\mathbb{P}(Z_n > 0|\Pi); U_n/V_n > c\big] \\ &\geq \mathbb{E}\big[U_n/V_n; U_n/V_n > c\big] \geq c\,\mathbb{P}(U_n/V_n > c)\ . \end{aligned} \tag{3.22}$$

The bound on the second factorial moment of Z, conditioned on Π, (3.19) and (3.22) imply[16]

$$\begin{aligned} \mathbb{E}[Z_k(Z_k-1)|U_k/V_k > c; Z_k > 0] &= \frac{\mathbb{E}\big[\mathbb{E}[Z_k(Z_k-1)|\Pi]; U_k/V_k > c\big]}{\mathbb{P}(Z_k > 0, U_k/V_k > c)} \\ &\leq c^{-2}\frac{\mathbb{E}[V_k^{-1}; U_k/V_k > c]}{c\mathbb{P}(U_n/V_n > c)} \\ &\leq c^{-3}\ . \end{aligned} \tag{3.23}$$

Therefore, conditioned on $\{U_k/V_k > c\}$, the second factorial moment of Z_k is bounded by a constant. By a variation of the Markov inequality for nonnegative, discrete random variables (see [Nav97]), for any $a > 1$,

$$\mathbb{P}(Z_k > a|U_k/V_k > c, Z_k > 0) \leq \frac{\mathbb{E}[Z_k(Z_k-1)|U_k/V_k > c, Z_k > 0]}{a(a-1)}\ .$$

Thus, by (3.23),

$$\mathbb{P}(1 \leq Z_k \leq a|U_k/V_k > c, Z_k > 0) \geq 1 - \frac{c^{-3}}{a(a-1)}\ . \tag{3.24}$$

Next we will treat the probability of the event $\{Z_n > 0, U_k/V_k > c\}$. For this, let

$$\begin{aligned} \mathcal{F}_k &:= \sigma(Q_1,\ldots,Q_k,Z_1,\ldots,Z_k) \\ \text{and} \qquad \Pi_{k+1,n} &:= \sigma(Q_{k+1},\ldots,Q_n) \end{aligned}$$

Note that Z_k, V_k and U_k are independent of $\Pi_{k+1,n}$. We use the fact that the mapping $x \to 1 - s^x$ is concave for $x \geq 0$, $s \in (0,1)$. Then by Jensen's inequality:

$$\begin{aligned} \mathbb{P}(Z_n > 0, &U_k/V_k > c) = \mathbb{E}\big[\mathbb{P}(Z_n > 0|\Pi); U_k/V_k > c\big] \\ &= \mathbb{E}\big[1 - f_{k+1,n}(0)^{Z_k}; U_k/V_k > c, Z_k > 0\big] \\ &= \mathbb{P}(U_k/V_k > c, Z_k > 0)\mathbb{E}\big[\mathbb{E}[1 - f_{k+1,n}(0)^{Z_k}|U_k/V_k > c, Z_k > 0, \Pi_{k+1,n}]\big|U_k/V_k > c, Z_k > 0\big] \\ &\leq \mathbb{E}\big[1 - f_{k+1,n}(0)^{\mathbb{E}[Z_k|U_k/V_k > c, Z_k > 0]}; U_k/V_k > c, Z_k > 0\big]\ . \end{aligned} \tag{3.25}$$

As $V_k^{-1} \geq 1$, $U_k/V_k > c$ implies $U_k^{-1} = e^{S_k} \leq c^{-1}$. Using (3.22), this yields

$$\begin{aligned} \mathbb{E}[Z_k|U_k/V_k > c, Z_k > 0] &= \frac{\mathbb{E}\big[\mathbb{E}[Z_k; Z_k > 0|\Pi]; U_k/V_k > c\big]}{\mathbb{P}(U_k/V_k > c, Z_k > 0)} \\ = \frac{\mathbb{E}\big[e^{S_k}; U_k/V_k > c\big]}{\mathbb{P}(U_k/V_k > c, Z_k > 0)} &\leq c^{-2}\ . \end{aligned}$$

Next we set $m := \lceil c^{-2} \rceil$ and use the fact that $1 - s^a$ is concave for $s \geq 0$ and $a \in \mathbb{N}$. By (3.25), as $\{U_k/V_k > c, Z_k > 0\}$ and $f_{k+1,n}$ are independent,

$$\begin{aligned} \mathbb{P}(Z_n > 0, U_k/V_k > c) &\leq \mathbb{E}[1 - f_{k+1,n}(0)^m; U_k/V_k > c, Z_k > 0] \\ &= \mathbb{P}(U_k/V_k > c, Z_k > 0)\mathbb{E}[1 - f_{k+1,n}(0)^m] \\ &\leq \mathbb{P}(U_k/V_k > c, Z_k > 0)(1 - \mathbb{P}(Z_{n-k-1} = 0)^m)\ . \end{aligned} \tag{3.26}$$

In the last step, Jensen's inequality was used again.

[16]by (3.14) $V_n^{-1} \leq 1$

Next we turn to the probability of having just one individual at time k. Conditioned on the environment,

$$\begin{aligned}\mathbb{P}(Z_{k+1}=1,Z_n>0|\Pi) &\geq \sum_{j=1}^{a}\mathbb{P}(Z_k=j,Z_{k+1}=1,Z_n>0|\Pi)\\ &\geq \mathbb{P}(1\leq Z_k\leq a|\Pi)\mathbb{P}(Z_{k+1}=0|\Pi,Z_k=1)^{a-1}\mathbb{P}(Z_{k+1}=1|\Pi,Z_k=1)\mathbb{P}(Z_n>0|\Pi,Z_{k+1}=1)\ .\end{aligned}$$

Each time when there is just one individual left, the process starts independently again. Note that Assumption 3.4 implies that there exists an $\epsilon_1>0$ such that $\mathbb{E}[Q(\{1\})Q(\{0\})]>\epsilon_1$. Thus, by Assumption 3.4 and with (3.24) for a large enough

$$\begin{aligned}\mathbb{P}(Z_{k+1}=1,U_k/V_k>c,Z_n>0) &\geq \mathbb{P}(1\leq Z_k\leq a,U_k/V_k>c)\\ &\quad \mathbb{E}\big[\mathbb{P}(Z_{k+1}=0|\Pi,Z_k=1)^{a-1}\mathbb{P}(Z_{k+1}=1|\Pi,Z_k=1)\big]\mathbb{E}\big[\mathbb{P}(Z_n>0|\Pi,Z_{k+1}=1)\big]\\ &\geq \mathbb{P}(U_k/V_k>c,Z_k>0)\ \mathbb{P}(1\leq Z_k\leq a|U_k/V_k>c,Z_k>0)\\ &\quad \mathbb{E}\big[\mathbb{P}(Z_{k+1}=0|\Pi,Z_k=1)^{a-1}\mathbb{P}(Z_{k+1}=1|\Pi,Z_k=1)\big]\ \mathbb{E}\big[\mathbb{P}(Z_n>0|\Pi,Z_{k+1}=1)\big]\\ &\geq p\mathbb{P}(U_k/V_k>c,Z_k>0)\mathbb{P}(Z_{n-k-1}>0)\end{aligned}$$

for some $p>0$. By (3.26), we end up with (w.l.o.g. $m>1$)

$$\begin{aligned}\mathbb{P}(Z_{k+1}=1|Z_n>0,U_k/V_k>c) &\geq p\,\frac{1-\mathbb{P}(Z_{n-k-1}=0)}{1-\mathbb{P}(Z_{n-k-1}=0)^m}\\ &\geq p\ m^{-1}\end{aligned}$$

which is the desired result. □

3.4.3 Proof of Theorem 3.1.3

Define

$$S^n \ := \ \Big(S_{\lfloor nt\rfloor}/a_n\Big)_{0\leq t\leq 1}\ . \tag{3.27}$$

In [AGKV05a, Lemma 2.3] it is proved that, under Assumption 3.5, for every $x\geq 0$

$$\mathcal{L}\big(S^n|L_n\geq -x\big) \ \xrightarrow{d}\ \mathcal{L}\big(L^+\big)\quad \text{as } n\to\infty\ ,$$

where L^+ is the meander of a strictly stable Lévy process.[17] For the case $x=0$, this fact has been proved in [BD94]. Analogously,

Lemma 3.4.5. *Under Assumption 3.5 and for $x\leq 0$,*

$$\mathcal{L}\big(S^n|M_n<-x\big) \ \xrightarrow{d}\ \mathcal{L}\big(L^-\big)\quad \textit{as } n\to\infty\ ,$$

where L^- is a strictly stable Lévy process conditioned to stay negative.

Let $\Phi:\mathcal{D}[0,1]\to\mathbb{R}_0^+$ be a bounded, continuous function and $\tilde{S}^n$ be the dual process of S^n. Then by Lemma 3.4.5

$$\mathbf{E}[\Phi(\tilde{S}^n)|M_n<-x] \ \to\ \mathbf{E}[\Phi(L^-)]\quad \text{as } n\to\infty$$

for every Φ and $x\geq 0$. For $g\in\mathcal{D}[0,1]$ and $\Phi:\mathcal{D}[0,1]\to\mathbb{R}_0^+$ bounded, continuous define

$$\tilde{\Phi}(g) \ := \ \Phi\Big(\big(g(1)-g((1-s)-)\big)_{0\leq s\leq 1}\Big)\ ,$$

where

$$g((1-s)-) \ := \ \lim_{\epsilon\searrow 0} g(1-s-\epsilon)\ .$$

[17] see Theorem 2.2.5

$\big(g(1)-g((1-s)-)\big)_{0\le s\le 1}$ is again a càdlàg-function and $\tilde\Phi : \mathcal{D}[0,1] \to \mathbb{R}_0^+$ is as Φ a bounded, continuous function. Thus, if a sequence of càdlàg functions $(g_n)_{n\in\mathbb{N}}$ converges to g_∞ with respect to the Skorohod topology, then

$$\tilde\Phi(g_n) \overset{n\to\infty}{\longrightarrow} \Phi\big(\big(g_\infty((1-s)-)\big)_{0\le s\le 1}\big) .$$

The dual version of Lemma 3.4.5 looks as follows:

Lemma 3.4.6. *Under Assumption 3.5,*

$$\mathcal{L}\big(S^n|\tau_n = n\big) \quad \overset{d}{\to} \quad \mathcal{L}\big(L^\dagger\big) \quad \text{as } n\to\infty ,$$

where $L^\dagger$ is the dual process of L^-. This is a strictly stable Lévy process conditioned to have its minimum at the end.

Now let $r\in\mathbb{N}$ be fixed. Define

$$S^{r,n} \quad := \quad \Big((S_{\lfloor r+t(n-r)\rfloor} - S_r)/a_n\Big)_{0\le t\le 1} .$$

By Lemma 3.4.6, for any $x \le 0$, $\mathcal{L}(S^{r,n}|M_n < -x)$ converges in distribution to the corresponding conditioned Lévy process $\mathcal{L}(L^-)$.

As in the proof of Theorem 3.1.1, a decomposition according to the minimum of $(S_n)_{n\in\mathbb{N}_0}$ is made. Let $m\in\mathbb{N}$. Then

$$\begin{aligned}
\mathbb{E}[\Phi(S^n)|Z_n>0] \;&=\; \mathbb{P}(Z_n>0)^{-1}\sum_{k=0}^{n-m}\mathbb{E}[\Phi(S^n);Z_n>0,\tau_n=k]\\
&+\;\mathbb{P}(Z_n>0)^{-1}\sum_{k=n-m}^{n}\mathbb{E}[\Phi(S^n);Z_n>0,\tau_n=k]\\
&=:\; s_1+s_2 .
\end{aligned}$$

By exactly the same arguments as in the proof of Theorem 3.1.1 and also using the result of Theorem 3.1.1, the first sum can be bounded by[18]

$$\limsup_{n\to\infty} s_1 \quad \le \quad c\sum_{k=m}^{\infty}k^{-1}a_k^{-1} \;<\; \infty$$

with some finite constant c. Taking $m\to\infty$, s_1 can be made abitrarily small.

In the following, the same definitions as on p. 33 are used. The second sum can be written as

$$\begin{aligned}
s_2 \;&=\; \mathbb{P}(Z_n>0)^{-1}\gamma^n\sum_{k=n-m}^{n}\mathbf{E}[e^{-S_n}\Phi(S^n)(1-f_{0,n}(0)),\tau_k=k,\min\{S_k,\dots S_n\}\ge S_k]\\
&=\; \mathbb{P}(Z_n>0)^{-1}\gamma^n\sum_{k=0}^{m}\mathbf{E}[e^{-\hat S_k}\mathbf{E}[e^{-S_{n-k}}\Phi(S^n)(1-f_{0,n-k}(\hat A_k))|\tau_{n-k}=n-k,\hat\Pi_k],\hat L_k\ge 0]\\
&\qquad\qquad\qquad\qquad\qquad\qquad\cdot\mathbf{P}(\tau_{n-k}=n-k) .
\end{aligned}\tag{3.28}$$

The main task of the proof is to separate the term connected with the survival probability, $e^{-S_{n-k}}(1-f_{0,n-k}(\hat A_k))$, from the part that converges to the Lévy process, S^n. For this purpose, instead of

$$\mathbf{E}\big[\Phi(S^n)e^{-S_{n-k}}(1-f_{0,n-k}(\hat A_k))\big|\tau_{n-k}=n-k,\hat\Pi_k\big] , \tag{3.29}$$

for fixed $r\in\mathbb{N}$,

$$\mathbf{E}\big[\Phi(S^{r,n})e^{-S_r}(1-f_{0,r}(\hat A_k))\big|\tau_n=n,\hat\Pi_k\big] \tag{3.30}$$

[18] recall that Φ is bounded

is considered. Later, it will be proved that the difference of these two expectations converges to 0 as $r \to \infty$.

First, the limit of (3.30) is treated. By duality,

$$\begin{aligned} \mathbf{E}[e^{-S_r}\Phi(S^{r,n})(1-f_{0,r}(s))|\tau_n = n] &= \mathbf{P}(\tau_n = n)^{-1}\mathbf{E}[\Phi(S^{r,n})e^{-S_r}(1-f_{0,r}(s)); \tau_n = n] \\ &= \mathbf{P}(M_n < 0)^{-1}\mathbf{E}[e^{-S_r}\tilde{\Phi}(S^{r,n})(1-f_{r,0}(s)); M_n < 0] . \end{aligned} \tag{3.31}$$

Define

$$\Pi_r := \sigma(Q_1, Q_2, \ldots, Q_r) . \tag{3.32}$$

Next, by adding $\wedge$, we again denote independent copies of the corresponding random variables. As $S^{r,n}$ is independent of Π_r,

$$\begin{aligned} &\mathbf{P}(M_n < 0)^{-1}\mathbf{E}[e^{-S_r}\tilde{\Phi}(S^{r,n})(1-f_{r,0}(s)); M_n < 0] \\ &\quad = \mathbf{P}(M_n < 0)^{-1}\mathbf{E}[\mathbf{E}[e^{-S_r}\tilde{\Phi}(S^{r,n})(1-f_{r,0}(s)); M_n < 0|\Pi_r]] \\ &\quad = \mathbf{P}(M_n < 0)^{-1}\mathbf{E}\big[e^{-\hat{S}_r}(1-\hat{f}_{r,0}(s))\mathbf{E}[\tilde{\Phi}(S^{r,n}); M_{n-r} < -\hat{S}_r|\hat{S}_r]; \hat{M}_r < 0\big] \\ &\quad = \mathbf{E}\left[\frac{\mathbf{P}(M_{n-r} < -\hat{S}_r|\hat{S}_r)}{\mathbf{P}(M_n < 0)}e^{-\hat{S}_r}(1-\hat{f}_{r,0}(s))\frac{\mathbf{E}[\tilde{\Phi}(S^{r,n}); M_{n-r} < -\hat{S}_r|\hat{S}_r]}{\mathbf{P}(M_{n-r} < -\hat{S}_r|\hat{S}_r)}; \hat{M}_r < 0\right] . \end{aligned}$$

By Lemma 3.4.5, as $n \to \infty$,

$$\frac{\mathbf{E}[\tilde{\Phi}(S^{r,n}); M_{n-r} < -\hat{S}_r|\hat{S}_r]}{\mathbf{P}(M_{n-r} < -\hat{S}_r|\hat{S}_r)} \to \mathbf{E}[\tilde{\Phi}(L^-)] \quad \text{a.s.}$$

where the limit does not depend on $\hat{S}_r$. Furthermore, as $n \to \infty$, for every $x \leq 0$, (see Lemma 3.2.1)

$$\frac{\mathbf{P}(M_{n-r} < -x)}{\mathbf{P}(M_n < 0)} \to v(x) .$$

Thus, recalling the definition of $\mathbf{E}^-$ on p. 25 (note that $\wedge$ can be dropped now),

$$\begin{aligned} &\lim_{n\to\infty} \mathbf{E}[e^{-S_r}\Phi(S^{r,n})(1-f_{0,r}(s))|\tau_n = n] \\ &\quad = \mathbf{E}[\tilde{\Phi}(L^-)]\mathbf{E}\big[v(S_r)e^{-S_r}(1-f_{r,0}(s)); M_r < 0\big] \\ &\quad = \mathbf{E}[\tilde{\Phi}(L^-)]\mathbf{E}^-\big[e^{-S_r}(1-f_{r,0}(s))\big] . \end{aligned} \tag{3.33}$$

Thus,

$$\mathbf{E}[e^{-S_r}\Phi(S^{r,n})(1-f_{0,r}(\hat{A}_k))|\tau_n = n, \hat{\Pi}_k] \to \mathbf{E}[\tilde{\Phi}(L^-)]\mathbf{E}^-\big[e^{-S_r}(1-f_{r,0}(\hat{A}_k))|\hat{\Pi}_k\big] \quad \text{a.s.} \tag{3.34}$$

Next, the difference of the expectations in (3.29) and (3.30) is examined. Define

$$\zeta_r^{(n)}(s) := \mathbf{E}[e^{-S_{n-k}}\Phi(S^n)(1-f_{0,n-k}(s))|\tau_{n-k} = n-k] - \mathbf{E}[e^{-S_r}\Phi(S^{r,n})(1-f_{0,r}(s))|\tau_n = n] .$$

As Φ is bounded, for $n \to \infty$,

$$\begin{aligned} |\zeta_r^{(n)}(s)| &= \big|\mathbf{E}[e^{-S_r}\Phi(S^{r,n})(1-f_{0,r}(s))|\tau_n = n] - \mathbf{E}[e^{-S_{n-k}}\Phi(S^n)(1-f_{0,n-k}(s))|\tau_{n-k} = n-k]\big| \\ &\leq \big|\mathbf{E}[e^{-S_r}\Phi(S^{r,n})(1-f_{0,r}(s))|\tau_n = n] - \mathbf{E}[e^{-S_r}\Phi(S^n)(1-f_{0,r}(s))|\tau_{n-k} = n-k]\big| \\ &\qquad + \big|\mathbf{E}[e^{-S_r}\Phi(S^n)(1-f_{0,r}(s))|\tau_{n-k} = n-k] - \mathbf{E}[e^{-S_{n-k}}\Phi(S^n)(1-f_{0,n-k}(s))|\tau_{n-k} = n-k]\big| \\ &=: \chi_1 + \chi_2 . \end{aligned} \tag{3.35}$$

As to the first term, recalling the definition of Π_r,

$$\chi_1 = \mathbf{P}(\tau_n = n)^{-1}\big|\mathbf{E}\big[e^{-S_r}(1-f_{0,r}(s))\big(\mathbf{E}[\Phi(S^{r,n}); \tau_n = n|\Pi_r] - \mathbf{E}[\Phi(S^n); \tau_{n-k} = n-k|\Pi_r]\big)\big]\big| .$$

By Lemma 3.4.6, $\mathbf{E}[\Phi(S^n)|\tau_{n-k}=n-k]$ and $\mathbf{E}[S^{r,n}|\tau_n=n]$ have the same limit. Thus, using Lemma 3.4.6 and dominated convergence (recall $e^{-S_r}(1-f_{0,r}(s))\le 1$ and Φ bounded)

$$\chi_1 \;\le\; \mathbf{E}\Big[e^{-S_r}(1-f_{0,r}(s))\frac{\mathbf{E}[\Phi(S^{r,n});\tau_n=n|\Pi_r]-\mathbf{E}[\Phi(S^n);\tau_{n-k}=n-k|\Pi_r]}{\mathbf{P}(\tau_n=n)}\Big] \;\overset{n\to\infty}{\longrightarrow}\; 0\ .$$

As discussed in the proof of Theorem 3.1.1, $e^{-S_n}(1-f_{n,0}(s))$ is monotone and bounded, hence converges $\mathbf{P}^-$–a.s. as $n\to\infty$ and has the limit $\mathcal{U}_\infty(s)$.

Thus, applying Proposition 3.2.3, the second term in (3.35) can be bounded by

$$\begin{aligned}\chi_2 \;&\le \mathbf{E}\Big[\sup|\Phi|\big|e^{-S_r}(1-f_{0,r}(s))-e^{-S_{n-k}}(1-f_{0,n-k}(s))\big|\Big|\tau_{n-k}=n-k\Big]\\ &\overset{n\to\infty}{\longrightarrow}\; \sup|\Phi|\cdot\mathbf{E}^-\big[|e^{-S_r}(1-f_{0,r}(s))-\mathcal{U}_\infty(s)|\big]\ .\end{aligned}$$

and thus

$$\limsup_{n\to\infty}|\zeta_r^{(n)}(s)| \;\le\; \sup|\Phi|\cdot\mathbf{E}^-\big[|e^{-S_r}(1-f_{0,r}(s))-\mathcal{U}_\infty(s)|\big]\ .$$

Therefore, using these results in (3.35), as $r\to\infty$ and uniformly in $s\in[0,1]$,

$$\mathbf{E}^-\big[|e^{-S_r}(1-f_{0,r}(s))-\mathcal{U}_\infty(s)|\big] \;\to\; 0\ . \tag{3.36}$$

Applying this and Theorem 3.1.1 to the decomposition (3.28),

$$\begin{aligned} s_2 =\;& \mathbb{P}(Z_n>0)^{-1}\gamma^n\sum_{k=0}^m\mathbf{E}\big[e^{-\hat S_k}\mathbf{E}[\Phi(S^n)e^{-S_{n-k}}(1-f_{0,n-k}(\hat A_k))|\tau_{n-k}=n-k,\hat\Pi_k];\hat L_k\ge 0\big]\mathbf{P}(\tau_{n-k}=n-k)\\ =\;& \mathbb{P}(Z_n>0)^{-1}\gamma^n\sum_{k=0}^m\mathbf{P}(\tau_{n-k}=n-k)\mathbf{E}\big[e^{-\hat S_k}\zeta_r^{(n)}(\hat A_k);\hat L_k\ge 0\big]\\ &+\mathbb{P}(Z_n>0)^{-1}\gamma^n\sum_{k=0}^m\mathbf{E}\big[e^{-\hat S_k}\mathbf{E}[\Phi(S^{r,n})e^{-S_r}(1-f_{0,r}(\hat A_k))|\tau_n=n,\hat\Pi_k];\hat L_k\ge 0\big]\mathbf{P}(\tau_{n-k}=n-k)\\ =:\;& s_3\;+\;s_4\ .\end{aligned}$$

In view of Theorem 3.1.1 and as $e^{-\hat S_k}\le 1$ for $\hat L_k\ge 0$, the first sum in the above equation can be bounded by

$$\begin{aligned}&\limsup_{n\to\infty}\mathbb{P}(Z_n>0)^{-1}\gamma^n\sum_{k=0}^m\mathbf{P}(\tau_{n-k}=n-k)\mathbf{E}\big[e^{-\hat S_k}\zeta_r^{(n)}(\hat A_k);\hat L_k\ge 0\big]\\ &\qquad\le\;\theta^{-1}\;m\;\sup|\Phi|\cdot\sup_{s\in[0,1]}\mathbf{E}^-\big[|e^{-S_r}(1-f_{0,r}(s))-\mathcal{U}_\infty(s)|\big]\ .\end{aligned}$$

By (3.36), taking the limit $r\to\infty$, the above sum vanishes, i.e. $s_3\to 0$.

By (3.34), the second sum above, s_4 tends to

$$s_4 \;\to\; \theta^{-1}\mathbf{E}[\tilde\Phi(L^-)]\sum_{k=0}^m\mathbf{E}\big[e^{-\hat S_k}e^{-S_r}(1-f_{0,r}(\hat A_k));\hat L_k\ge 0\big]$$

as $n\to\infty$. Taking the limit $r\to\infty$ and applying dominated convergence yields

$$s_4 \;\to\; \theta^{-1}\mathbf{E}[\tilde\Phi(L^-)]\sum_{k=0}^m\mathbf{E}\big[e^{-\hat S_k}\mathbf{E}^-[\mathcal{U}_\infty(\hat A_k)|\Pi_k];\hat L_k\ge 0\big]\ .$$

Finally letting $m\to\infty$, s_1 tends to zero. Thus by (3.18) and dropping the $\wedge$,

$$\begin{aligned} s_4 \;&\to\; \theta^{-1}\sum_{k=0}^\infty\mathbf{E}[\tilde\Phi(L^-)]\mathbf{E}\Big[\mathbf{E}^-[\mathcal{U}_\infty(A_k)|\Pi_k\big]e^{-S_k};L_k\ge 0\Big]\\ &=\;\mathbf{E}[\tilde\Phi(L^-)]\ ,\end{aligned}$$

which proves the theorem. □

Chapter 4

Large deviations

The second part of this dissertation deals with ***large deviations*** of BPREs. The main problem studied here is illustrated by the following example:

Figure 4.1: Common hawthorn (taken from *http://en.wikipedia.org/wiki/File:Common_hawthorn.jpg*).

Consider the toy model of a population of apomictic plants with a one year life cycle, touched in the introduction. The reproductive success of these plants depends on the environmental conditions, changing from one year to the other in an i.i.d. manner. Say one apomictic individual (e.g. a Crataegus, commonly called hawthorn or thornapple) is brought into a new region to develop itself. After a long time, an observer comes back. Now imagine the observer has all necessary information about the distribution of the i.i.d. random environment (i.e. the distribution of X and some properties of the offspring distribution Q which will be specified in the sequel). Say, the new region is unfavorable for the individuals and the population is expected to be extinct. Instead, the obverser finds a very large population. This may be due to extraordinary large offspring numbers within an 'ordinary' environment, 'ordinary' reproduction within an unusually favourable environment or a mixture of both. In other words, what is the least improbable way of greatly deviating from the expected population size? Or, to speak in terms of a gambler: what is the optimal strategy to attain an extraordinarily large population size? This question will be formalized in the following sections in terms of the rate function which describes the exponential decay rate of the large deviation probabilities. Following the terminology in [dH00], if for some measurable set A,

$$\mathbb{P}(Z_n \in A) \quad = \quad e^{-an+o(n)} \ ,$$

then a will be referred to as ***cost for*** A. Minimizing the costs for all possible paths attaining large values then yields the proper rate function for the BPRE.

The main focus of this chapter is studying upper deviations. First, limit theorems under the law $\mathbb{P}$ are formulated, i.e. probabilities averaged over the environment are studied (so-called *annealed approach*). In Section 4.4, lower deviations of BPREs are studied for offspring distributions with linear fractional generating functions.

The asymptotics of the large deviation probabilities, conditioned on the environment (*quenched approach*), which are substantially different, will be studied in Section 4.5.

Recently, upper large deviations of BPREs under the annealed approach have been studied in [Koz06] for geometric offspring distributions and in [BB09] for general offspring distributions. In the latter paper, only a lower bound for the (upper) large deviation probabilities is obtained which is improved here. In the particular case of geometric offspring distributions, direct calculations of generating functions are feasible and the asymptotic of the (upper) large deviation probabilities (including lower order terms) is obtained in [Koz06]. In this chapter, the rate function found in [Koz06] for geometric offspring distributions is generalized and the second order phase transition touched in [Koz06] is explained in detail.

Some related results for the lower deviations of BPREs in special cases (e.g. branching processes without extinction) can also be found in [BB09]. Recently, large deviations of BPREs have also been studied in [HL10].

As it will be proved in the sequel, for (upper) large deviations of BPREs, there are phase transitions of order two (i.e. discontinuities of the second derivative of the rate function), corresponding to structural changes of the least costly paths. In Section 4.2, offspring distributions with geometrically bounded tails are studied. In this case, the probability that an individual has exponentially many children is of a lower order than exponential and, as it will be specified later, large deviations are essentially realized in a 'good' environment. The case of general offspring distributions, where ***jumps*** (i.e. one individual having exponentially many children) have only exponentially small propability, is treated in Section 4.3. The main Theorem 4.2.2 from Section 4.2 will also be covered by the general Theorem 4.3.2 in Section 4.3. Anyhow, Theorem 4.2.2 will also be proved in detail, as some auxiliary results for BPREs with offspring distributions with geometrically bounded tails are of independent interest, and as the proofs in Section 4.2 are technically less involved.

By $\mathbb{P}_z(\cdot)$ we denote that the process starts with $Z_0 = z$-many individuals. Unless otherwise specified, the initial population size is one, i.e. $Z_0 \equiv 1$.

4.1 Preliminaries

A first large deviation statement is readily obtained: The limit

$$\gamma \quad := \quad \lim_{n\to\infty} -\frac{1}{n} \log \mathbb{P}(Z_n > 0) \tag{4.1}$$

exists and $0 \leq \gamma < \infty$ (up to the degenerated case $\mathbb{P}(Z_1 = 0) = 1$ which we exclude in the sequel). Moreover,

$$\mathbb{P}(Z_n > 0) \quad \leq \quad e^{-\gamma n} \tag{4.2}$$

for all n. This follows from

$$\mathbb{P}(Z_{n+m} > 0) \quad \geq \quad \mathbb{P}(Z_n > 0)\, \mathbb{P}(Z_m > 0)\ .$$

Thus the sequence $(-\log \mathbb{P}(Z_n > 0))_n$ is subadditive, and (4.1) results from properties of nonnegative subadditive sequences (see [DZ93]).

As explained in the preceding sections, the environment is essentially described by the associated random walk, $(S_n)_{n\geq 0}$ with initial state $S_0 = 0$. Recall its definition,

$$S_n \quad = \quad X_1 + \ldots + X_n\ , \quad n \geq 1\ ,$$

where the increments are given by the logarithmic mean offspring number in each generation,

$$X_k \quad = \quad \log \sum_{y=0}^{\infty} y\ Q_k(\{y\}), \quad k \geq 1\ .$$

To describe the cost for an extraordinarily good environment (characterized by the associated random walk), we require the rate function $\Lambda : \mathbb{R} \to \bar{\mathbb{R}}^+$ which fulfills for all $\theta \in \mathbb{R}$

$$\begin{aligned} \limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}(S_n \geq \theta n) \quad &\leq \quad -\inf_{y\geq\theta} \Lambda(y) \;=\; \Lambda(\theta)\ , \\ \liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}(S_n > \theta n) \quad &\geq \quad -\inf_{y>\theta} \Lambda(y) \;=\; \Lambda(\theta+)\ , \end{aligned}$$

and

$$\mathbb{P}(S_n \geq \theta n) \quad \leq \quad e^{-\Lambda(\theta) n} \quad . \tag{4.3}$$

It is a convex, nondecreasing function, given by

$$\Lambda(\theta) \quad = \quad \sup_{s\geq 0} \left\{ s\theta - \log \varphi(s) \right\} \tag{4.4}$$

with $\varphi(s) = \mathbb{E}[e^{sX}]$. As Λ is convex and lower semicontinuous, there is at most one $\theta \geq 0$ with $\Lambda(\theta) \neq \Lambda(\theta+)$. In this case, $\Lambda(\theta+) = \infty$ (see e.g [dH00], [DZ93]).

Remark. *Usually, Λ is defined as the Legendre transform of* $\log \varphi$ *and the supremum in (4.4) is taken over all $s \in \mathbb{R}$. Here, we are only interested in upper deviations, thus setting $\Lambda(\theta) = 0$ for $\theta \leq \mathbb{E}[X]$ is convenient.*

Throughout this chapter, we need existence of a proper rate function for the associated random walk. This is assured by the following assumption:

Assumption 4.1. *There is a $s > 0$ such that the moment generating function*

$$\varphi(s) \quad = \quad \mathbb{E}\left[e^{sX}\right] \;<\; \infty\ .$$

In particular $\mathbb{E}[X] \geq -\infty$ exists.

4.2 Offspring distributions with geometrically bounded tails

Here, large deviations for BPREs are studied under the assumption that the offspring distributions have geometrically bounded tails (Assumption 4.2). As explained at the beginning of this chapter, the main result in [Koz06] is generalized here. The main part of this section is published in [BK10].

Assumption 4.2. *There are constants $k_0 \in \mathbb{N}_0$, $0 \leq a < b$ and $c > 0$ such that Q a.s. takes values in the set of all probability distributions $\mathcal{A} \subset \Delta$ with the following property:*
If R has distribution $\mathcal{P}$ and expectation $\mathcal{E}[R] = m$, then

$$\mathcal{E}\left[(R-j)^+\right] \quad \leq \quad c\ m\ \left(\frac{a+m}{b+m}\right)^{j-k_0}\ , \quad j \geq k_0\ . \tag{4.5}$$

Note that $\mathcal{E}[(R-j)^+]$ is decreasing in j. It is required that this takes place at a geometric rate, where the rate may slow down as $\mathcal{E}[R]$ gets larger. Degenerated cases (0 carrying most of the mass) are excluded by the linear factor m in (4.5). Essentially, Assumption 4.2 assures that R has geometrically bounded tails. This is illustrated by the following examples:

Examples. • *geometric distributions with success probability p and expectation $\frac{1-p}{p}$: $\mathcal{P}(R \geq i) = (1-p)^i = (\frac{m}{m+1})^i$ and*

$$\mathcal{E}\left[(R-j)^+\right] \quad = \quad \sum_{i>j} \mathcal{P}(R \geq i) \quad = \quad m \left(\frac{m}{m+1}\right)^j$$

fulfills (4.5) for $k_0 = 0$, $a = 0$, $b = 1$, and $c = 1$.

• *distributions fulfilling the following condition:*
There are constants $c > 0$ and $d \in (0,1)$ such that

$$\mathcal{P}(R = j) \quad \leq \quad c\ m\ d^j \quad . \tag{4.6}$$

In contrast to (4.5), (4.6) implies the geometric decay rate and thus also the expectations are uniformly bounded by a constant.

• *distributions with Gaussian tails:*
There are constants $\alpha, c > 0$ such that

$$\mathcal{P}(R = j) \quad \leq \quad c\ m\ \exp(-\alpha j^2) \quad .$$

• *distributions with support in $\{0, \dots, k_0\}$ trivially fulfill condition (4.5).*

As turns out, under Assumption 4.2, (upper) large deviations of Z are determined by a convex function Γ_γ only depending on the ***cost of survival***, γ (see (4.1)) and the rate function for S, Λ (see (4.4)).

Recall $\gamma \geq 0$. Γ_γ is then defined as the largest convex function fulfilling

$$\Gamma_\gamma(0) \leq \gamma\ , \qquad \Gamma_\gamma(\theta) \leq \Lambda(\theta)$$

for all $\theta \geq 0$. It is not difficult to see that this function is given by

$$\Gamma_\gamma(\theta) \quad = \quad \begin{cases} \gamma\left(1 - \frac{\theta}{\theta^*}\right) + \frac{\theta}{\theta^*}\Lambda(\theta^*) & , \text{ if } \theta < \theta^* \\ \Lambda(\theta) & , \text{ else} \end{cases} \tag{4.7}$$

where $0 \leq \theta^* \leq \infty$ is such that

$$\frac{\Lambda(\theta^*) - \gamma}{\theta^*} \quad = \quad \inf_{\theta \geq 0} \frac{\Lambda(\theta) - \gamma}{\theta} \ . \tag{4.8}$$

As will be proved in Section 4.2.2, Γ is also represented by:

Lemma 4.2.1. *For any $\theta \geq 0$*

$$\Gamma_\gamma(\theta) = \inf_{0<t\leq 1} \left\{t\gamma + (1-t)\Lambda(\theta/(1-t))\right\} \ .$$

Figure 4.2 illustrates Γ_γ in the case $\gamma < \Lambda(0)$. If $\gamma \geq \Lambda(0)$ then $\theta^* = 0$ and $\Gamma_\gamma = \Lambda$.

The following theorem has been obtained in [Koz06] in the case of geometric offspring distributions.

Theorem 4.2.2. *Let Assumptions 4.1 and 4.2 be fulfilled. Then for any $\theta \geq 0$*

$$\begin{aligned} \limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}(Z_n \geq e^{\theta n}) \quad &\leq \quad -\Gamma(\theta) \quad , \\ \liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}(Z_n > e^{\theta n}) \quad &\geq \quad -\Gamma(\theta+) \end{aligned}$$

where $\Gamma = \Gamma_\gamma$.

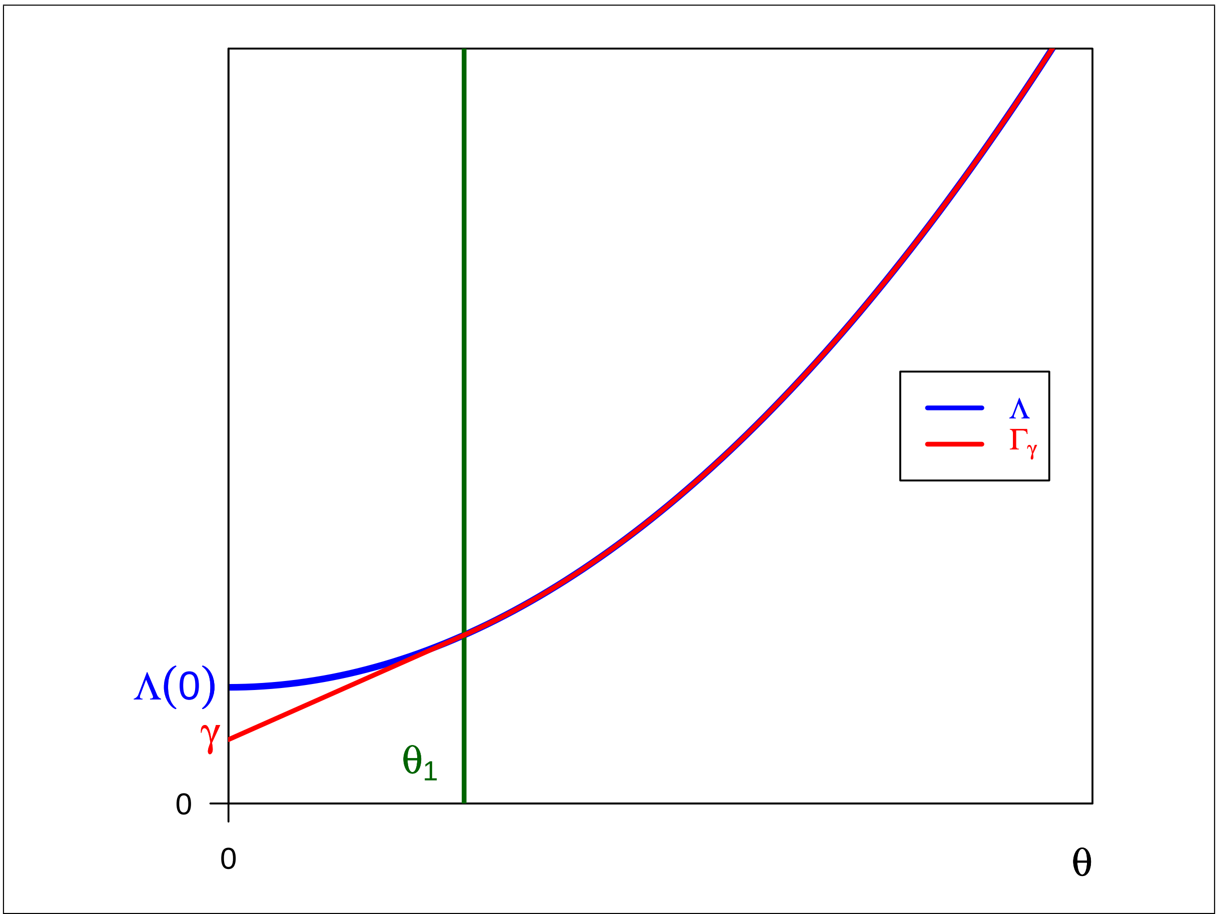


Figure 4.2: Γ_γ and Λ in the case of $\gamma < \Lambda(0)$ with $\theta_1 = \theta^*$.

Theorem 4.2.2 says that a phase transition of order two occurs if $\gamma < \Lambda(0)$. Then $\theta^* > 0$ and there are different regimes $\theta < \theta^*$ and $\theta \geq \theta^*$. Recall the classification of BPREs from Chapter 1. In the supercritical and critical case, $\gamma = \Lambda(0) = 0$, whereas in the weakly and intermediately subcritical case, $\gamma = -\log \mathbb{E}[e^{\beta X}]$. By definition of β, (see (1.5)),

$$\mathbb{E}[e^{\beta X}] \quad = \quad \inf_{s\geq 0} \mathbb{E}[e^{sX}]$$

and thus

$$\log \mathbb{E}[e^{\beta X}] \quad = \quad \sup_{s\geq 0} \mathbb{E}[e^{sX}] \quad = \quad \Lambda(0) \ .$$

Therefore, under the mild assumptions explained in Chapter 2 (see also [Afa80], [BGK05]), $\gamma = \Lambda(0)$ also for the weakly and intermediately subcritical case. $\gamma < \Lambda(0)$ occurs only for strongly subcritical BPREs (i.e. $\mathbb{E}[m(Q)\log(m(Q))] < 0$, see [GKV03]). Indeed in this case, $\mathbb{E}[X\exp(X)] < 0$, thus the differentiation of $s \to \mathbb{E}[\exp(sX)]$ in $s = 1$ is negative and $\Lambda(0) = \sup_{s\geq 0}\{-\log(\mathbb{E}[\exp(sX)])\} > -\log(\mathbb{E}[\exp(X)]) = \gamma$.

Interpretation of Theorem 4.2.2

By Theorem 4.2.2, the rate function Γ for the BPRE only depends on the exponential decay rate of the survival probability γ and Λ, the rate function of S. Under the assumption of geometrically bounded tails (see Assumption 4.2), the fine structure of the offspring distributions is not of importance.

This results from the fact that the large deviation event $\{Z_n \geq e^{\theta n}\}$ is essentially realized in an exceptional environment and not by exceptionally big offspring numbers. This would require either exponentially many individuals reproducing in an exceptional manner or one individual having exponentially many offsprings. Both probabilities are (by Assumption 4.2 for the latter) of lower order than exponential.

Next, recall that

$$\mathbb{P}(S_n \geq 0) = e^{-\Lambda(0)n+o(n)} \quad , \quad \mathbb{P}(Z_n \geq 1) = e^{-\gamma n+o(n)} \ .$$

In the case of $\gamma = \Lambda(0)$, it is natural to expect that the events $\{S_n \geq \theta n\}$ and $\{Z_n \geq e^{\theta n}\}$ essentially coincide as then the cost for keeping the associated random walk nonnegative is the same as the cost of survival. From our theorem we see that in the case of $\gamma < \Lambda(0)$ this is also true if $\theta \geq \theta^*$. For $\theta < \theta^*$, however, matters change. There we also have to consider the events $\{Z_{tn} \geq 1, S_n - S_{tn} \geq \theta n\}$ with $0 < t < 1$, which in view of (4.1) have exponentially small probabilities as well. Surprisingly, for t properly chosen, this event has exponentially larger probability than $\{S_n \geq \theta n\}$. Thus it is of advantage to keep the population just alive at the beginning and to enforce exponential growth only later. For the environment, this means that S first decreases linearly up to time tn and then increases linearly. For further details we refer to the proof in Section 4.2.2.

For a better understanding of the meaning of Γ, note that Γ_0 determines the large deviations of $S_n - L_n$, where[19]

$$L_n \quad = \quad \min\{S_0, \dots, S_n\} \ .$$

That is, under Assumption 4.1 and for $\theta \geq 0$,

$$\begin{aligned} \limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(S_n - L_n \geq \theta n\right) &\leq -\Gamma_0(\theta) \ , \\ \liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(S_n - L_n > \theta n\right) &\geq -\Gamma_0(\theta+) \ . \end{aligned}$$

This result immediately follows from the proof of Theorem 4.2.2 (see Section 4.2.2): As we shall see, for $\{Z_n \geq e^{\theta n}\}$, the occurence of $\{S_n \geq \theta n\}$ is not required, but essentially only that of $\{S_n - L_n \geq \theta n\}$ and survival of Z until L_n is attained. The latter event has exponentially small probability, represented by γ.

As touched before, the rate function is essentially the minimum of the cost of all possible ways of attaining $\{Z_n \geq e^{\theta n}\}$, n being large. By the representation (4.7), for $\theta < \theta^*$, the best strategy is a period of survival of length $(1 - \theta/\theta^*)t$, succeeded by a period of geometric growth (S growing linearly) of length $\theta/\theta^* n$. Thus, for $\theta < \theta^*$, the linear part has always the same slope, namely θ^*. To see this, note that the infimum in the representation of Γ in Lemma 4.2.1 is reached at t_θ such that $\theta/(1 - t_\theta) = \theta^*$. We will now show that the infimum is taken for $t = t_\theta$. Calculating the derivative with respect to t yields that the condition for the infimum,

$$\begin{aligned} 0 &= \gamma - \Lambda(\theta/(1 - t_\theta)) + \Lambda'(\theta/(1 - t_\theta))\theta/(1 - t_\theta) \\ &= \gamma - \Lambda(\theta^*) + \theta^* \Lambda'(\theta^*) \ . \end{aligned}$$

By (4.8), the above equation is fulfilled. This means that smaller values of θ correspond to a longer period where there is just survival of the process.

The most probable paths are illustrated in Figure 4.3, where the generation time and the logarithmic generation sizes have been scaled by n.

For a related result in the context of polymers see [dH00, section IX.9].

In the next section, we discuss two characteristics of the distribution of Z_n, which are of independent interest. Theorem 4.2.2 is proved in Section 4.2.2.

4.2.1 Two characteristics of Z

As we will see in this section, Assumption 4.2 assures that certain characteristics of the distribution of Z_n, known for the linear fractional case, are useful also in the more general case treated here. We derive bounds for the normalized variance and tail probabilities in terms of the associated random walk.

[19] Here, we consider the infimum over $0 \leq k \leq n$ instead of $1 \leq k \leq n$ as in the definition in Chapter 1.

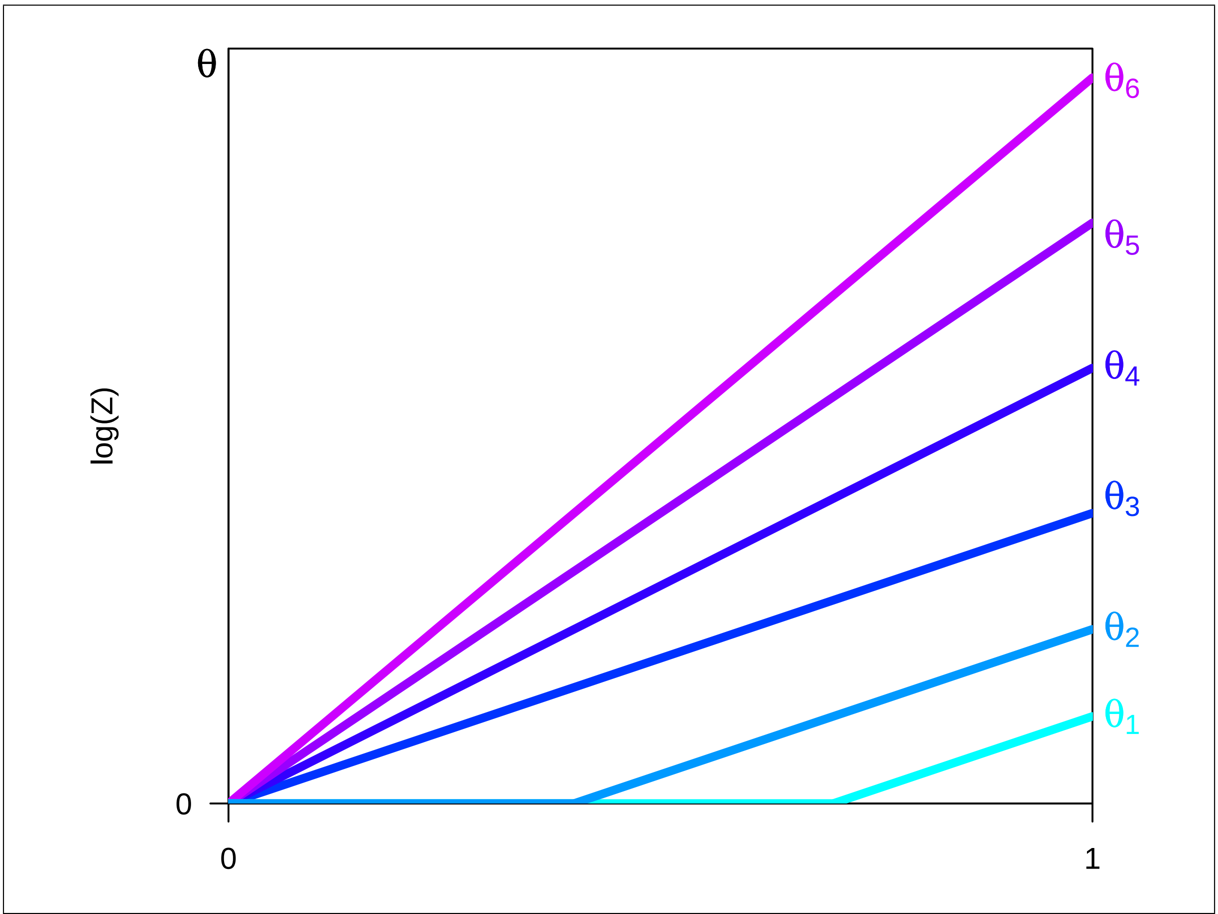


Figure 4.3: Most probable paths in the strongly subcritical case ($\gamma < \Lambda(0)$) for different θ with $\theta_1 < \theta_2 < \theta_3 = \theta^* < \theta_4 < \theta_5 < \theta_6$.

These bounds are finer than a bound for the tail probability derived in Section 4.3 for general offspring distributions and are of independent interest. Let

$$U_n := e^{-S_n}$$
$$V_n := \sum_{k=0}^{n} e^{-S_k}.$$

Then $U_n = \mathbb{E}[Z_n|\Pi]^{-1}$. The following two results shed some light on the significance of V_n.

Proposition 4.2.3. *Under Assumption 4.2, there is an $\alpha < \infty$ such that*

$$\frac{\mathbb{E}[Z_n^2|\Pi]}{\mathbb{E}[Z_n|\Pi]^2} \leq \alpha V_n \quad a.s.$$

Remark. *Proposition 4.2.3 holds under much weaker conditions than Assumption 4.2. Namely, it is only required that for some constant $0 < c < \infty$,*

$$\frac{\sum_{k=0}^{\infty} k(k-1)Q(k)}{\left(\sum_{k=0}^{\infty} kQ(k)\right)^2} \leq \frac{c}{m(Q)} \quad a.s. \tag{4.9}$$

E. g. it suffices that the standardized second factorial moment of the offspring distributions is uniformly bounded by a constant.

By Proposition 4.2.3, the BPRE, conditioned on the environment, has a standardized variance of at most order $O(ne^{-L_n})$. Assume S_n being large and the associated random walk having a low minimum (L_n very small). Then with a high probability (of order e^{L_n}), the BPRE becomes extinct. With a small probability, the BPRE survives and then will be very large (of order $e^{S_n-L_n}$). Thus the standardized variance will be large. On the other hand, if S_n is large and $L_n \sim 0$, then the standarized variance of Z_n, conditioned on the environment, is of constant order.

For the tail probabilities, the following estimate holds:

Theorem 4.2.4. *Under Assumption 4.2, there is a $\beta > 0$ such that for all $z \geq 0$:*

$$\mathbb{P}(Z_n \geq z|\Pi) \quad \leq \quad 2 \ \exp\left(-\frac{\beta U_n}{V_n}\ z\right) \quad a.s.$$

or likewise

$$\mathbb{P}\left(\frac{Z_n}{\mathbb{E}[Z_n|\Pi]} \geq z \middle| \Pi\right) \quad \leq \quad 2\exp\left(-\frac{\beta}{V_n}\ z\right) \quad a.s.$$

For the proof, we use a formula derived in [GK00]. Let

$$f_n(s) = \sum_{k=0}^{\infty} s^k Q_n(\{k\})$$

be the probability generating function of the offspring distribution of an individual in generation $n-1$. Note that $X_n = \log f_n'(1)$. Recall (3.11)

$$\mathbb{E}\big[s^{Z_n}|\Pi\big] \quad = \quad f_1(f_2(\cdots f_n(s)\cdots)) \ = \ f_{0,n}(s), \quad s \geq 0$$

and the definition from (3.12),

$$\begin{aligned} f_{k,n} \quad &= \quad f_{k+1} \circ f_{k+2} \circ \cdots \circ f_n, \quad 0 \leq k < n; \ f_{n,n} = id \\ g_k(s) \quad &= \quad \frac{1}{1-f_k(s)} - \frac{1}{f_k'(1)(1-s)}, \quad s \geq 0. \end{aligned}$$

Below, we shall see that the singularity of $g_k(s)$ at $s = 1$ is removable under Assumption 4.2. We have

$$U_k \quad = \quad (f_1'(1)\cdots f_k'(1))^{-1} \ = \ f_{0,k}'(1)^{-1} \ .$$

We use the formula derived in Section 3.3, (3.13):

$$\frac{1}{1-f_{0,n}(s)} \quad = \quad \frac{U_n}{1-s} + \sum_{k=0}^{n-1} U_k g_{k+1}(f_{k+1,n}(s)), \quad s \geq 0.$$

Note that the above relation does not only hold for those s in the domain of convergence of $f_{0,n}(s)$, but for all $s \geq 0$. A statement similar to the following lemma can be found in [GK00] for $s \in [0, 1)$:

Lemma 4.2.5. *Let $f(s) = \sum_{k=0}^{\infty} s^k p_k$ be the generating function of R with distribution $q = (p_k)$ and expectation $\mathcal{E}[R] = m$. Then under Assumption 4.2, the function*

$$h(s) \quad = \quad \frac{1}{1-f(s)} - \frac{1}{m(1-s)}$$

is continuous everywhere, and there is a number $d < \infty$ such that for all $s \geq 0$

$$0 \quad \leq \quad h(s) \quad \leq \quad d\left(\frac{1}{m}+1\right).$$

In particular $\mathcal{E}[R(R-1)] \leq d(m+m^2)$.

Proof. Defining $r_j = \sum_{k>j} p_k$, we rewrite $f(s)-1$ to extract the factor $s-1$ (see [Fel68, chapter XI]):

$$f(s)-1 \quad = \quad (s-1)\sum_{j=0}^{\infty} s^j r_j \;=\; (s-1)g(s) \quad ,$$

where

$$g(s) \quad = \quad \sum_{j=0}^{\infty} s^j r_j \quad . \tag{4.10}$$

In the same manner

$$\begin{aligned} g(s)-g(1) \quad &= \quad (s-1)\sum_{j=0}^{\infty} s^j (r_{j+1}+r_{j+2}+\cdots) \\ &= \quad (s-1)k(s) \quad , \end{aligned}$$

with

$$k(s) \quad = \quad \sum_{j=0}^{\infty} s^j (r_{j+1}+r_{j+2}+\cdots) \quad .$$

Note that by (4.10), $g(1)=m$ and thus

$$\begin{aligned} \frac{1}{1-f(s)} - \frac{1}{m(1-s)} \quad &= \quad \frac{1}{m(s-1)} - \frac{1}{(s-1)g(s)} \\ &= \quad \frac{k(s)}{mg(s)} \quad . \end{aligned}$$

Since $\mathcal{E}[(\xi-j)^+] = r_{j+1}+r_{j+2}+\ldots$, by Assumption 4.2 the functions $f(s)$, $g(s)$ and $k(s)$ are finite for $s<\frac{b+m}{a+m}$. Therefore g is continuous in $s=1$ and thus everywhere.
Next let $1 \le s < \frac{1}{2}+\frac{b+m}{2(a+m)}$, implying $s \le \frac{b}{a}$. As h is nondecreasing, $g(s)\ge m$ for all $s\ge 1$. Thus (also recall $m = r_0 + r_1 + \ldots$),

$$\begin{aligned} \frac{1}{1-f(s)} - \frac{1}{m(1-s)} \quad &\le \quad \frac{k(s)}{m^2} \\ &= \quad \frac{1}{m}\sum_{j=0}^{\infty} s^j \frac{r_{j+1}+r_{j+2}+\cdots}{r_0+r_1+\cdots} \\ &\le \quad \frac{1}{m}\left(\sum_{j=0}^{k_0-1} s^j + c s^{k_0}\sum_{j=0}^{\infty} s^j \left(\frac{a+m}{b+m}\right)^j\right) \\ &\le \quad \frac{k_0\left(\frac{b}{a}\right)^{k_0}}{m} + \frac{2\left(\frac{b}{a}\right)^{k_0} c(b+m)}{(b-a)m} \quad , \end{aligned} \tag{4.11}$$

where (4.5) has been used in the prelast step. The last inequality follows from $s \le \frac{b}{a}$ in the first sum and $s < \frac{1}{2}\frac{b+m}{2(a+m)}$ in the second geometric summation.
For $s \ge \frac{1}{2}+\frac{b+m}{2(a+m)}$, the negative term is dropped:

$$\begin{aligned} -\frac{1}{f(s)-1} + \frac{1}{m(s-1)} \quad &\le \quad \frac{1}{m(s-1)} \\ &\le \quad \frac{2(a+m)}{m(b-a)} \quad . \end{aligned} \tag{4.12}$$

Note that the last estimate does not require the assumption $f(s)<\infty$. Choosing a sufficiently large constant d, (4.11) and (4.12) prove the claim for every $s\ge 1$. For $s<1$, $h(s)\le 2h(1)$ for $0\le s\le 1$ is used (see [GK00, Lemma 2.1]). The last claim results from $h(1) = m^{-2}\mathcal{E}[R(R-1)]$. Note that for the last claim, only condition (4.9) is required. □

Proof of Proposition 4.2.3. For the proof of the proposition, we need an expression for $f''_{0,n}(1)$, which is well-known (see e.g. [Afa93]). From $f_{0,n} = f_{0,n-1} \circ f_n$, by chain rule for differentiation $f'_{0,n}(1) = f'_{0,n}(1) f'_n(1)$ and $f''_{0,n}(1) = f''_{0,n-1}(1)(f'_n(1))^2 + f'_{0,n-1}(1) f''_n(1)$, thus

$$\frac{f''_{0,n}(1)}{(f'_{0,n}(1))^2} \quad = \quad \frac{f''_{0,n-1}(1)}{(f'_{0,n-1}(1))^2} + \frac{f''_n(1)}{f'_{0,n-1}(1)(f'_n(1))^2} \quad .$$

Recalling $U_k = e^{-S_k}$ and $V_n = \sum_{k=0}^n e^{-S_k}$, the last estimate in Lemma 4.2.5 implies

$$\frac{f''_n(1)}{f'_{0,n-1}(1)(f'_n(1))^2} \quad \le \quad d(U_{n-1} + U_n)$$

and thereby

$$\frac{\mathbb{E}[Z_n(Z_n-1)|\Pi]}{\mathbb{E}[Z_n|\Pi]^2} \quad = \quad \frac{f''_{0,n}(1)}{(f'_{0,n}(1))^2} \;\le\; 2dV_n \quad \text{a.s.}$$

As $V_n \ge U_n = (\mathbb{E}[Z_n|\Pi])^{-1}$, choosing $\alpha = 2d+1$ yields the claim. □

Proof of Theorem 4.2.4. We obtain from (3.13) and Lemma 4.2.5

$$\frac{1}{1 - f_{0,n}(s)} \quad \le \quad \frac{U_n}{1-s} + d\sum_{k=0}^{n-1}(U_k + U_{k+1}) \;\le\; \frac{U_n}{1-s} + 2dV_n \; . \tag{4.13}$$

As $f_{0,n}(s) > 1$ for $s > 1$, we only have a useful bound if the right-hand side of (4.13) is negative. Thus for $s < 1 + \frac{U_n}{2dV_n}$, we get

$$f_{0,n}(s) \quad \le \quad \frac{U_n - (2dV_n - 1)(s-1)}{U_n - 2dV_n(s-1)} \; . \tag{4.14}$$

For $s \le 1 + \frac{U_n}{4dV_n}$, since the right-hand side in (4.14) is nondecreasing for $s \ge 1$, and as $V_n \ge 1$,

$$f_{0,n}(s) \quad \le \quad 1 + \frac{1}{2d} \; .$$

Therefore (w.l.o.g. $d \ge \frac{1}{2}$), for every $1 \le s \le 1 + \frac{U_n}{4dV_n}$,

$$\mathbb{P}(Z_n \ge z|\Pi) \quad \le \quad s^{-z} \; f_{0,n}(s) \;\le\; 2s^{-z} \quad \text{a.s.} \tag{4.15}$$

Note that $\frac{U_n}{V_n} \le 1$, so there is a $\beta > 0$ such that $e^{\beta\frac{U_n}{V_n}} \le 1 + \frac{U_n}{4dV_n}$ (β is defined by $e^\beta = 1 + \frac{1}{4d}$). Taking $s = e^{\beta\frac{U_n}{V_n}}$ in (4.15) yields Theorem 4.2.4. □

4.2.2 Proof of Theorem 4.2.2

Here we prove the representation of Γ from Lemma 4.2.1, which is used in the sequel: For any $\theta \ge 0$

$$\Gamma(\theta) = \inf_{0<t\le 1}\left\{t\gamma + (1-t)\Lambda(\theta/(1-t))\right\} .$$

Proof of Lemma 4.2.1. Let us denote this infimum by $\iota(\theta)$. We show that it fulfills the properties defining $\Gamma(\theta)$. First for any $\theta', \theta'' \ge 0$, $\lambda \in (0,1)$ and $\epsilon > 0$ there are $t', t'' \in (0,1]$ such that, applying convexity of Λ,

$$\begin{aligned}
\lambda\iota(\theta') + (1-\lambda)\iota(\theta'') \;&\ge\; \lambda t'\gamma + \lambda(1-t')\Lambda(\theta'/(1-t')) + (1-\lambda)t''\gamma + (1-\lambda)(1-t'')\Lambda(\theta''/1-t'')) - \epsilon \\
&= \Big(\lambda(1-t') + (1-\lambda)(1-t'')\Big)\gamma + (\lambda t' + (1-\lambda)(1-t''))\frac{\lambda(1-t')}{\lambda(1-t') + (1-\lambda)(1-t'')}\Lambda(\theta'/(1-t')) \\
&\quad + (\lambda(1-t') + (1-\lambda)(1-t''))\frac{(1-\lambda)(1-t'')}{\lambda(1-t') + (1-\lambda)(1-t'')}\Lambda(\theta''/(1-t'')) - \epsilon \\
&\ge \Big(1 - (\lambda(1-t') + (1-\lambda)(1-t''))\Big)\gamma + (\lambda(1-t') + (1-\lambda)(1-t''))\Lambda\Big(\frac{\lambda\theta' + (1-\lambda)\theta''}{\lambda(1-t') + (1-\lambda)(1-t'')}\Big) - \epsilon \\
&\ge \iota(\lambda\theta' + (1-\lambda)\theta'') - \epsilon \; .
\end{aligned}$$

Letting $\epsilon \to 0$ yields the convexity of ι. Next choosing $t = 1$ implies $\iota(\theta) \leq \Lambda(\theta)$ and letting $t \to 0$ entails $\iota(0) \leq \gamma$. Finally let $\kappa(\theta)$ be any convex function below $\Gamma(\theta)$ and γ. Then for any $t \in (0, 1]$, $\theta \geq 0$

$$\begin{aligned} t\gamma + (1-t)\Lambda(\theta/(1-t)) &\geq t\kappa(0) + (1-t)\kappa(\theta/(1-t)) \\ &\geq \kappa\big(t0 + \lambda(\theta/(1-t))\big) \\ &= \kappa(\theta) \ . \end{aligned}$$

It follows $\iota(\theta) \geq \kappa(\theta)$, and the proof is complete. □

The upper bound

We follow ideas of Kozlov.

Lemma 4.2.6. *Under Assumptions 4.1 and 4.2, for any $\theta \geq 0$:*

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}(Z_n \geq e^{\theta n}) \quad \leq \quad -\inf_{y \geq \theta} \Gamma(y).$$

Proof. We restrict ourselves to $\theta > 0$ as the case $\theta = 0$ is covered by (4.1). Let

$$L_n \quad = \quad \min_{0 \leq k \leq n} S_k$$

and $\epsilon > 0$ such that $\theta - \epsilon > 0$. We have

$$\begin{aligned} \mathbb{P}(Z_n \geq e^{\theta n}) &= \mathbb{P}(Z_n \geq e^{\theta n}, S_n - L_n \geq (\theta-\epsilon)n) + \mathbb{P}(Z_n \geq e^{\theta n}, S_n - L_n < (\theta-\epsilon)n) \\ &=: p_{1n} + p_{2n} \quad . \end{aligned}$$

Now, as Z_k is independent of $S_n - S_k$ and by (4.2) and (4.3),

$$\begin{aligned} \mathbb{P}(Z_n \geq e^{\theta n}, S_n - L_n \geq (\theta-\epsilon)n) &\leq \sum_{k=0}^{n-1} \mathbb{P}(Z_k > 0, S_n - S_k \geq (\theta-\epsilon)n) \\ &\leq \sum_{k=0}^{n-1} e^{-\gamma k} e^{-\Lambda\left((\theta-\epsilon)\frac{n}{n-k}\right)\cdot(n-k)} \\ &= \sum_{k=0}^{n-1} \exp\left(-n\left(\gamma\frac{k}{n} + \Lambda\left((\theta-\epsilon)\frac{n}{n-k}\right)\frac{n-k}{n}\right)\right) \ . \end{aligned}$$

In view of Lemma 4.2.1

$$p_{1n} \quad \leq \quad n e^{-\Gamma(\theta-\epsilon)n} \ .$$

As to p_{2n}, by means of Theorem 4.2.4,

$$\mathbb{P}(Z_n \geq e^{\theta n} | \Pi) \quad \leq \quad 2\exp\left(-\frac{\beta U_n}{V_n} e^{\theta n}\right) \ .$$

Now $V_n \leq (n+1)e^{-L_n}$, thus

$$\begin{aligned} p_{2n} &\leq \mathbb{E}\left[2\exp\left(-\frac{\beta U_n}{V_n}e^{\theta n}\right); S_n - L_n < (\theta-\epsilon)n\right] \\ &\leq \mathbb{E}\left[2\exp\left(-\beta(n+1)^{-1}e^{\theta n - (S_n - L_n)}\right); S_n - L_n < (\theta-\epsilon)n\right] \\ &\leq 2\exp\left(-\beta(n+1)^{-1}e^{\epsilon n}\right). \end{aligned}$$

By standard arguments from large deviation theory (see [dH00]), we have

$$\begin{aligned} \limsup_{n\to\infty} \frac{1}{n}\log \mathbb{P}(Z_n \geq e^{\theta n}) &\leq \limsup_{n\to\infty} \frac{1}{n} \log\left[p_{1n} + p_{2n}\right] \\ &= \max\left[\limsup_{n\to\infty} \frac{1}{n}\log p_{1n}, \limsup_{n\to\infty}\frac{1}{n}\log p_{2n}\right] \\ &\leq \max\left[-\Gamma(\theta-\epsilon), -\infty\right] \ . \end{aligned}$$

As Γ is left-continuous, taking the limit $\epsilon \to 0$ yields the result. □

The lower bound

Lemma 4.2.7. *Under Assumptions 4.1 and 4.2, for any $\theta \geq 0$:*

$$\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}(Z_n > e^{\theta n}) \quad \geq \quad -\inf_{y>\theta} \Gamma(y) \ .$$

A weaker result has been proved in [BB09] under different assumptions.

Proof. Without loss of generality, we restrict ourselves to the case $\Gamma(\theta+) < \infty$. For every $0 < t \leq 1$, by Markov property,

$$\begin{aligned} \mathbb{P}(Z_n > e^{\theta n}) &= \mathbb{P}(Z_{\lceil tn \rceil} > 0)\mathbb{P}(Z_n > e^{\theta n} \mid Z_{\lceil tn \rceil} > 0) \\ &\geq \mathbb{P}(Z_{\lceil tn \rceil} > 0)\mathbb{P}(Z_{\lfloor (1-t)n \rfloor} > e^{\theta n}) \ . \end{aligned}$$

We fix θ', θ'' with $\theta < \theta' < \theta''$. Then

$$\mathbb{P}(Z_{\lfloor (1-t)n \rfloor} > e^{\theta n}) \quad \geq \quad \mathbb{P}(Z_{\lfloor (1-t)n \rfloor} > e^{\theta n}, \theta' n < S_{\lfloor (1-t)n \rfloor} < \theta'' n) \ .$$

An inequality due to Paley and Zygmund (see e.g. [Kal01]) yields for $0 < r < 1$

$$\mathbb{P}(Z_{\lfloor (1-t)n \rfloor} > r\mathbb{E}[Z_{\lfloor (1-t)n \rfloor}|\Pi]|\Pi) \quad \geq \quad (1-r)^2 \frac{\mathbb{E}[Z_{\lfloor (1-t)n \rfloor}|\Pi]^2}{\mathbb{E}\left[Z^2_{\lfloor (1-t)n \rfloor}|\Pi\right]} \ .$$

From Proposition 4.2.3

$$\frac{\mathbb{E}[Z_{\lfloor (1-t)n \rfloor}|\Pi]^2}{\mathbb{E}\left[Z^2_{\lfloor (1-t)n \rfloor}|\Pi\right]} \quad \geq \quad \frac{1}{\alpha V_{\lfloor (1-t)n \rfloor}} \quad \geq \quad \frac{e^{L_{\lfloor (1-t)n \rfloor}}}{\alpha(n+1)}.$$

Thus with $r = e^{-(\theta'-\theta)} \geq e^{-(\theta'-\theta)n}$,

$$\begin{aligned} &\mathbb{P}(Z_{\lfloor (1-t)n \rfloor} > e^{\theta n}, \theta' n < S_{\lfloor (1-t)n \rfloor} < \theta'' n) \\ &\qquad = \mathbb{E}\left[\mathbb{P}(Z_{\lfloor (1-t)n \rfloor} > e^{\theta n}|\Pi); \theta' n < S_{\lfloor (1-t)n \rfloor} < \theta'' n)\right] \\ &\qquad \geq (\alpha(n+1))^{-1}(1-r)^2\mathbb{E}[e^{L_{\lfloor (1-t)n \rfloor}}; \theta' n < S_{\lfloor (1-t)n \rfloor} < \theta'' n] \\ &\qquad \geq (\alpha(n+1))^{-1}(1-r)^2\mathbb{P}(L_{\lfloor (1-t)n \rfloor} \geq 0, \theta' n < S_{\lfloor (1-t)n \rfloor} < \theta'' n) \ . \end{aligned}$$

Let $\tilde{\theta} = \frac{1}{2(1-t)}(\theta' + \theta'')$. First we assume $\varphi(s) < \infty$ for all $s \in \mathbb{R}^+$. Then $\Lambda(\tilde{\theta}) = \tau\tilde{\theta} - \log\varphi(\tau)$ defines τ (see e.g. [DZ93] for properties of the rate function) and we can change measure according to

$$\widetilde{\mathbb{P}}(X \in dx) \quad = \quad \rho^{-1}e^{\tau x}\mathbb{P}(X \in dx)$$

where $\rho = \varphi(\tau)$. Thus

$$\begin{aligned} &\mathbb{P}(L_{\lfloor (1-t)n \rfloor} \geq 0, \theta' n < S_{\lfloor (1-t)n \rfloor} < \theta'' n) \\ &\quad \geq \rho^{\lfloor (1-t)n \rfloor}\widetilde{\mathbb{E}}\left[e^{-\tau S_{\lfloor (1-t)n \rfloor}}; \theta'(1-t)^{-1}\lceil (1-t)n \rceil < S_{\lfloor (1-t)n \rfloor} < \theta''(1-t)^{-1}\lfloor (1-t)n \rfloor, L_{\lfloor (1-t)n \rfloor} \geq 0\right] \\ &\quad \geq \rho^{\lfloor (1-t)n \rfloor}e^{-\tau\theta''(1-t)^{-1}\lfloor (1-t)n \rfloor}\widetilde{\mathbb{P}}(\theta'(1-t)^{-1}\lceil (1-t)n \rceil < S_{\lfloor (1-t)n \rfloor} < \theta''(1-t)^{-1}\lfloor (1-t)n \rfloor, L_{\lfloor (1-t)n \rfloor} \geq 0) \ . \end{aligned}$$

S is under $\widetilde{\mathbb{P}}$ a random walk with $\widetilde{\mathbb{E}}[S_n] = \theta n$. Therefore $\widetilde{\mathbb{P}}(\theta'(1-t)^{-1}n < S_n < \theta''(1-t)^{-1}n) \to 1$ and $\widetilde{\mathbb{P}}(L_n \geq 0) \to \widetilde{\mathbb{P}}(L_\infty \geq 0) = p > 0$. Thus for n large enough,

$$\widetilde{\mathbb{P}}(\theta'(1-t)^{-1}\lceil (1-t)n \rceil < S_{\lfloor (1-t)n \rfloor} < \theta''(1-t)^{-1}\lfloor (1-t)n \rfloor, L_{\lfloor (1-t)n \rfloor} \geq 0) \geq \frac{p}{2}$$

and therefore, for any $0 < t \leq 1$

$$\begin{aligned} \liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}(Z_n > e^{\theta n}) \ &\geq \ \liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}(Z_{\lceil tn \rceil} > 0) \\ &\quad + \liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}(Z_{\lfloor (1-t)n \rfloor} > e^{\theta n}, \theta'(1-t)^{-1}\lceil (1-t)n \rceil < S_{\lfloor (1-t)n \rfloor} < \theta''(1-t)^{-1}\lfloor (1-t)n \rfloor) \\ &= -\left(t\gamma + (1-t)(\theta''(1-t)^{-1}\tau - \log\varphi(\tau))\right) . \end{aligned}$$

Letting $\theta', \theta'' \to \theta$ (as $\Lambda(\theta+) < \infty$, Λ is continuous in θ)

$$\theta''(1-t)^{-1}\tau - \log\varphi(\tau) \to \Lambda((1-t)^{-1}\theta+) \ .$$

For the general case of $\varphi(s) = \infty$, we condition on $\{\max_{i=1,\dots,\lfloor(1-t)n\rfloor} X_i < x\}$. For the conditioned random walk, the moment generating function is finite on $\mathbb{R}^+$ and we can find a $\tilde{\tau}_x$ such that the calculation above holds. Now

$$\begin{aligned}&\mathbb{P}(L_{\lfloor(1-t)n\rfloor} \geq 0, \theta' n < S_{\lfloor(1-t)n\rfloor} < \theta'' n)\\ &\geq \ \mathbb{P}\Big(L_{\lfloor(1-t)n\rfloor} \geq 0, \theta' n < S_{\lfloor(1-t)n\rfloor} < \theta'' n \Big| \max_{i=1,\dots,\lfloor(1-t)n\rfloor} X_i < x\Big) \ \mathbb{P}(X < x)^{\lfloor(1-t)n\rfloor}\end{aligned}$$

and the moment generating function of the conditioned random walk is $\mathbb{P}(X<x)^{-n}\varphi_x^n(s)$, where $\varphi_x(s) = \mathbb{E}[e^{sx}; X < x]$. Thus, letting $\theta', \theta'' \to \theta$ (and thereby $\tilde{\tau}_x \to \tau_x$),

$$\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}(Z_n > e^{\theta n}) \quad \geq \quad -\big(t\gamma + (1-t)(\theta(1-t)^{-1}\tau_x - \log\varphi_x(\tau_x))\big) \ ,$$

where

$$\begin{aligned}\theta(1-t)^{-1}\tau_x - \log\varphi_x(\tau_x) \quad &= \quad \theta(1-t)^{-1}\tau_x - \log\mathbb{E}[e^{\tau_x X}|X<x] + \log\mathbb{P}(X<x)\\ &= \quad \sup_{s\geq 0}\Big\{\theta(1-t)^{-1}s - \log\mathbb{E}[e^{sX}; X<x]\Big\} \ .\end{aligned}$$

The right-hand side is nonincreasing in x. Thus, by monotone convergence, we may interchange the limit $x \to \infty$ and the supremum and letting $x \to \infty$,

$$\begin{aligned}\theta(1-t)^{-1}\tau_x - \log\varphi_x(\tau_x)) \quad &\to \quad \sup_{s\geq 0}\Big\{\theta(1-t)^{-1}s - \log\mathbb{E}[e^{sX}]\Big\}\\ &= \quad \Lambda(\theta(1-t)^{-1}) \ .\end{aligned}$$

By Lemma 4.2.1,

$$\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}(Z_n > e^{\theta n}) \quad \geq \quad -\Gamma(\theta+) \ ,$$

which entails the result. □

4.3 Heavy-tailed offspring distributions

This section is based on joint work with Vincent Bansaye[20] and published in [BB10]. As in the preceeding section, the existence of a proper rate function for the associated random walk is required. Hence, throughout this whole section, let Assumption 4.1 be fulfilled.

We write $R = R(f)$ for a random variable associated with the probability generating function f:

$$\mathcal{E}[s^R] = f(s) \quad (0 \leq s \leq 1)$$

and we denote by $m = m(R) = m(f)$ its expectation:

$$m = f'(1) \ .$$

4.3.1 Main results and interpretation

As discussed in the preceding section, in the case of geometrically bounded tails, the rate function of Z only depends on properties of the associated random walk and not on the fine structure of the distributions $(Q_1, Q_2, \dots)$. However, if the offspring distributions may have heavy (i.e. polynomial) tails, matters change. Then one individual having exponentially many offsprings has only exponentially small

[20] École Polytechnique, CMAP, Palaiseau

probability and large deviations of Z_n may be realized by one individual having exceptionally many children. In the sequel, it will be shown that the rate function (denoted by ψ) indeed depends on the polynomial decay rate of the offspring distributions.

The following assumption ensures that the tail of the offspring distribution of an individual decays at least with exponent $\beta \in (1,\infty)$ (uniformly with respect to the environment).

Assumption $\mathcal{H}(\beta)$. *There is a constant $0 < d < \infty$ such that Q a.s. takes values in the set of all probability distributions $\mathcal{A} \subset \Delta$ with the following property:*
If R has distribution $\mathcal{P}$ and expectation $\mathcal{E}[R] = m$, then for all $z > 0$

$$\mathcal{P}(R > z \mid R > 0) \quad \leq \quad d\ (m \wedge 1)\ z^{-\beta} \quad a.s.$$

The rate function ψ of Z depends on β, γ and Λ and is defined by

$$\psi(\theta) \;=\; \psi_{\gamma,\beta,\Lambda}(\theta) \;:=\; \inf_{t\in[0,1],s\in[0,\theta]} \Big\{ t\gamma + \beta s + (1-t)\Lambda((\theta-s)/(1-t)) \Big\} \quad . \tag{4.16}$$

If $\gamma = 0$, that is in the supercritical ($\mathbb{E}[X] > 0$) and critical ($\mathbb{E}[X] = 0$) case , ψ can be expressed as

$$\psi(\theta) = \inf_{s\in[0,\theta]} \big\{ \beta s + \Lambda(\theta - s) \big\} \ .$$

The main result of this section establishes the rate function ψ of Z in the case of heavy-tailed offspring distributions. The first assumption in the following theorem assures that there are some offspring distributions with polynomial tails with exponent $-\beta$, and by $\mathcal{H}(\beta)$, no tail distribution exceeds this exponent.

Theorem 4.3.1. *Assume that for some $\beta \in (1,\infty)$, $\log(\mathbb{P}(Z_1 > z))/\log(z) \overset{z\to\infty}{\longrightarrow} -\beta$ and that additionally Assumption $\mathcal{H}(\beta)$ holds. Then for every $\theta \geq 0$,*

$$\frac{1}{n}\log(\mathbb{P}(Z_n \geq e^{\theta n})) \overset{n\to\infty}{\longrightarrow} -\psi(\theta) \ .$$

Remark. *This Theorem still holds if we just assume that there exists a slowly varying function l such that*

$$\mathcal{P}(R > z | R > 0) \quad \leq \quad d\ (m \wedge 1)\ l(z) z^{-\beta} \quad a.s. \tag{4.17}$$

instead of Assumption $\mathcal{H}(\beta)$. Indeed, by properties of slowly varying functions (see appendix, Section A.2), for any $\epsilon > 0$, there exists a constant d_ϵ such that $\mathcal{P}(R > z \mid R > 0) \leq d_\epsilon\ (m \wedge 1) z^{-\beta+\epsilon}$ a.s. As for fixed $\theta \geq 0$, $\psi_{\gamma,\beta,\Lambda}$ is continuous in β, letting $\epsilon \to 0$ yields the upper bound. Let us note that we can weaken Assumption $\mathcal{H}(\beta)$ by letting d depend on the environment. But this would make the proof more cumbersome. For the proof of the lower bound in Theorem 4.2.2, it is only required that $\mathbb{E}[\sum_{k=1}^{\infty} k^s \mathcal{P}(R > z \mid R > 0)] < \infty$ (see p. 60) for some $s > 1$ which is also assured by (4.17).

Let us consider two consequences of this result. First, a large deviation result for offspring distributions without heavy tails can be derived by letting $\beta \to \infty$, which generalizes Theorem 4.2.2. The following result is proved in Section 4.3.5 as a corollary from Theorem 4.3.1.

Theorem 4.3.2. *If Assumption $\mathcal{H}(\beta)$ is fulfilled for every $\beta > 0$, then for every $\theta \geq 0$,*

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}(Z_n \geq e^{\theta n}) \quad \leq \quad -\Gamma(\theta) \quad ,$$
$$\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}(Z_n > e^{\theta n}) \quad \geq \quad -\Gamma(\theta+) \ ,$$

where $\Gamma(\theta) = \inf_{t\in[0,1]}\{t\gamma + (1-t)\Lambda(\theta/(1-t))\}$.

For example, this result holds if $\mathcal{P}(R > z|R > 0) \leq c(m \wedge 1)\exp(-z^b)$ a.s. for some constant $b > 0$.

As second consequence, Theorem 4.3.1 can also be applied to the Galton-Watson case. Then the environment is not random and the offspring distribution q (with mean m) is deterministic, meaning $\Lambda(\theta) = \infty$

for $\theta > \log m$ and $\Lambda(\log m) = 0$. We refer to [BB93, Pro00] for precise results for large deviations without heavy tails. For the decay rate of the survival probability, it is known that (see [AN72]) in the subcritical case ($m < 1$)

$$\gamma \quad = \quad -\log m$$

and $\gamma = 0$ in the critical ($m = 1$) and supercritical ($m > 1$) case. Thus, in the subcritical case,

$$\psi(\theta) \quad = \quad \gamma t + \beta s \ .$$

In the critical and supercritical case, it remains to minimize

$$\psi(\theta) = \inf_{s \in [0,\theta]} \{\beta s + \Lambda(\theta - s)\} \ ,$$

where $\Lambda(\theta) = 0$ for $\theta \leq \log m$ and $\Lambda(\theta) = \infty$ for $\theta > \log m$. Hence,

$$\psi(\theta) = \beta(\theta - \log m) \ .$$

The following sections are organized as follows. In the next paragraphs, different characterizations of the rate function ψ are presented and the phase transitions are described. The lower bound in Theorem 4.3.1 is proved in Section 4.3.2, the upper bound in Sections 4.3.3 and 4.3.4 by distinguishing the case $\beta \in (1, 2]$ and the case $\beta > 2$. The proof for $\beta > 2$ is technically more involved since it requires higher order derivatives of generating functions. In Section 4.3.4 it is explained how to adapt the arguments of the proof for $\beta \in (1, 2]$ for the case $\beta > 2$.

Path interpretation of the rate function.

As precised by our theorem, the rate function yields the exponential decay rate of the probability of reaching exceptionally large values, namely

$$\mathbb{P}(Z_n \geq \theta n) \quad = \quad e^{-\psi(\theta)n + o(n)}.$$

A reasonable way to reach extraordinary large values, $\{Z_n \geq \exp(\theta n)\}$ for $n \gg 1$ and $\theta \geq \mathbb{E}[X]$, can be described as follows (see Figure 4.4). At the beginning, up to time $\lfloor tn \rfloor$, the process just survives. The probability of this event decreases as $e^{-\gamma \lfloor tn \rfloor}$. At time $\lfloor tn \rfloor$, there are very few individuals and one individual has extraordinarily many offsprings, namely e^{sn}-many. The probability of this event is given by $\mathbb{P}(Z_1 \geq \exp(sn))$, thus it is of the order of $e^{-\beta sn}$. Then the process grows geometrically according to its expectation in a good environment to reach $e^{\theta n}$. That is S grows linearly such that $S_n - S_{\lfloor nt \rfloor} \approx [\theta - s]n$. As results from Cramer's Theorem, the probability to observe this exceptionally good environment decreases as $e^{-(1-t)\Lambda((\theta-s)/(1-t))n}$. The most probable path is then to follow the way which minimizes the sum of these three *costs* $t\gamma$, βs and $(1 - t)\Lambda((\theta - s)/(1 - t))$ which yields the rate function ψ.

Thus the optimal strategy to realize large deviations is described by couples (t_θ, s_θ) such that

$$\psi(\theta) = t_\theta \gamma + \beta s_\theta + (1 - t_\theta)\psi((\theta - s_\theta)/(1 - t_\theta)) \ .$$

As detailed in the next paragraph, different strategies may occur following the regime of the process and the value of θ. In any case, if there is a *jump* ($s_\theta > 0$), then it occurs either at the beginning ($t_\theta = 0$) or at the end ($t_\theta = 1$) of the trajectory.

Obviously, keeping the population size small during a first period ($t_\theta > 0$) and then growth within a good environment can be relevant only in the subcritical case. Actually we see below that this situation can occur only if the process is strongly subcritical, as already discussed in Section 4.2 in the case of offspring distributions without heavy tails. In all subcritical cases, depending on β and Λ, the optimal stratetgy may be to keep the population size small by just surviving until the final time and then jump to the final value. This means that the $e^{\theta n}$-many individuals have a common ancestor in one of the last generations. The phase transitions will be precised later.

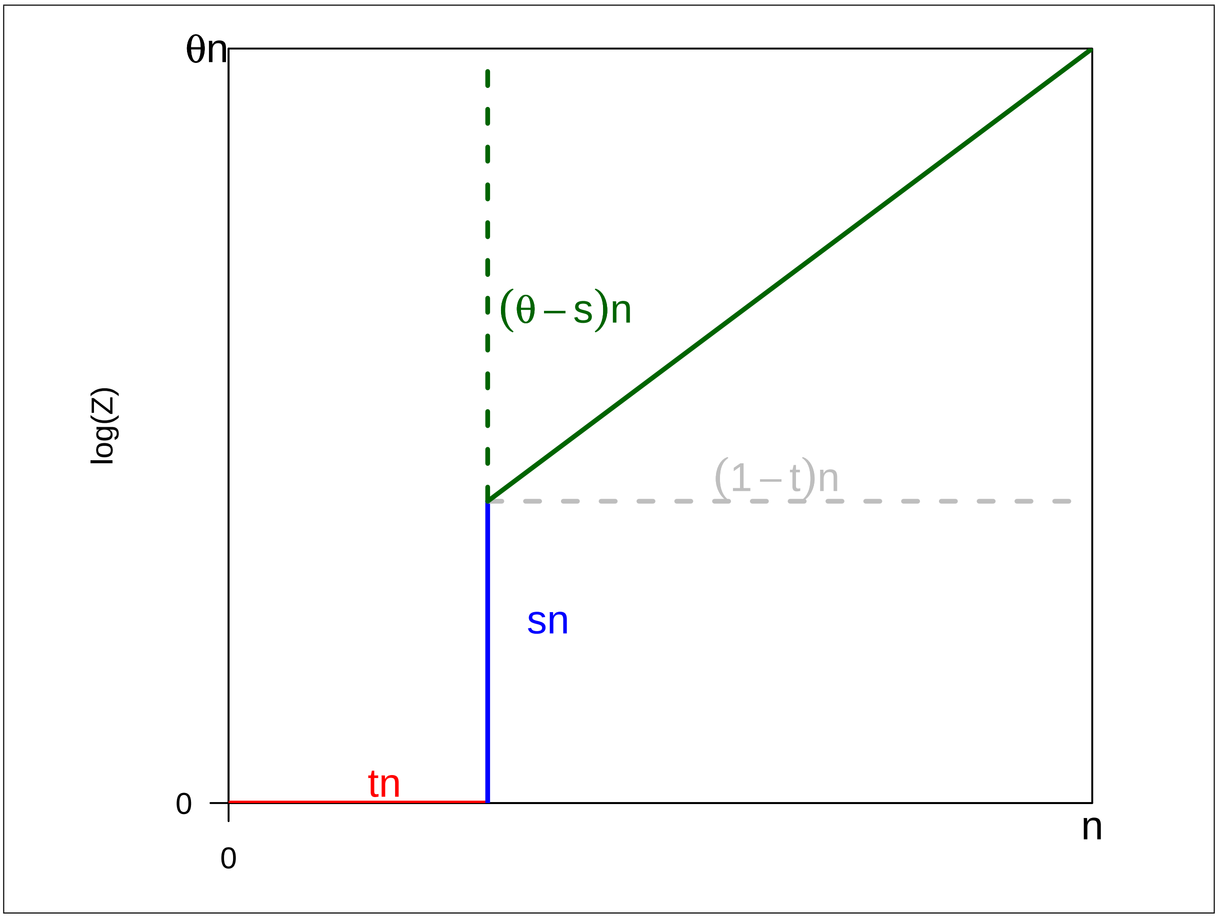


Figure 4.4: (Schematic) Representation of a possible path of $\log Z$ (to $\log Z_n = \theta n$) with a survival period and a jump, followed by linear growth.

In the supercritical case ($\mathbb{E}[X] > 0$), the process starts growing (geometrically) right from the beginning ($t_\theta = 0$). Or it may have a jump at time $t_\theta = 0$, and then grow geometrically.

More formally, following the proof of [BB09], it should be possible to prove the uniqueness of (t_θ, s_θ) (except for degenerated situations) and the forthcoming trajectorial result. We refrain from writing down more details here, as the proof becomes very heavy and technical.

Open Problem. *Prove that, conditioned on $Z_n \leq e^{cn}$,*

$$\sup_{t\in[0,1]} \left\{ \left| \log(Z_{[tn]})/n - f_\theta(t) \right| \right\} \stackrel{n\to\infty}{\longrightarrow} 0$$

in probability in the sense of the uniform norm where

$$f_\theta(t) := \begin{cases} 0 & , \quad \text{if } t \leq t_\theta \\ \beta s_\theta + \frac{c}{1-t_\theta}(t - t_\theta) & , \quad \text{if } t > t_\theta \ . \end{cases}$$

In the Galton-Watson case with mean offspring number m, the optimal strategy is either to survive until the final time and jump to the desired value θ (critical and subcritical case, $m \leq 1$), or to jump to $\theta - \log(m)$ and then grow according to the expectation (supercritical case, $m > 1$).

Graphical interpretation of the rate function.

In this section, another characterization of ψ is discussed. As proved in Lemma 4.3.8 (see Section 4.3.6), ψ is the largest convex function fulfilling for all $x, \theta \geq 0$

$$\psi(0) = \gamma, \quad \psi(\theta + x) \leq \psi(\theta) + \beta x, \quad \psi(\theta) \leq \Lambda(\theta) \ .$$

As explained in the preceeding section, the first condition will only play a role in the subcritical cases.

We define below the number $\theta^* \in [0,1]$ such that for every $\theta < \theta^*$, the process begins with a survival period (i.e. $t_\theta > 0$). Then the rate function ψ is linear for $\theta \in [0, \theta^*]$.

Similarly, we define $\theta^\dagger$ such that for $\theta > \theta^\dagger$ the strategy begins with a jump and the rate function is linear with slope β. In the subcritical cases, the best strategy may consist of just surviving until almost the end and then a jump to the terminal value, which corresponds to $\theta^\dagger = 0$, hence $\psi(\theta) = \gamma + \beta\theta$.

Figure 4.5 illustrates ψ in the strongly subcritical case.

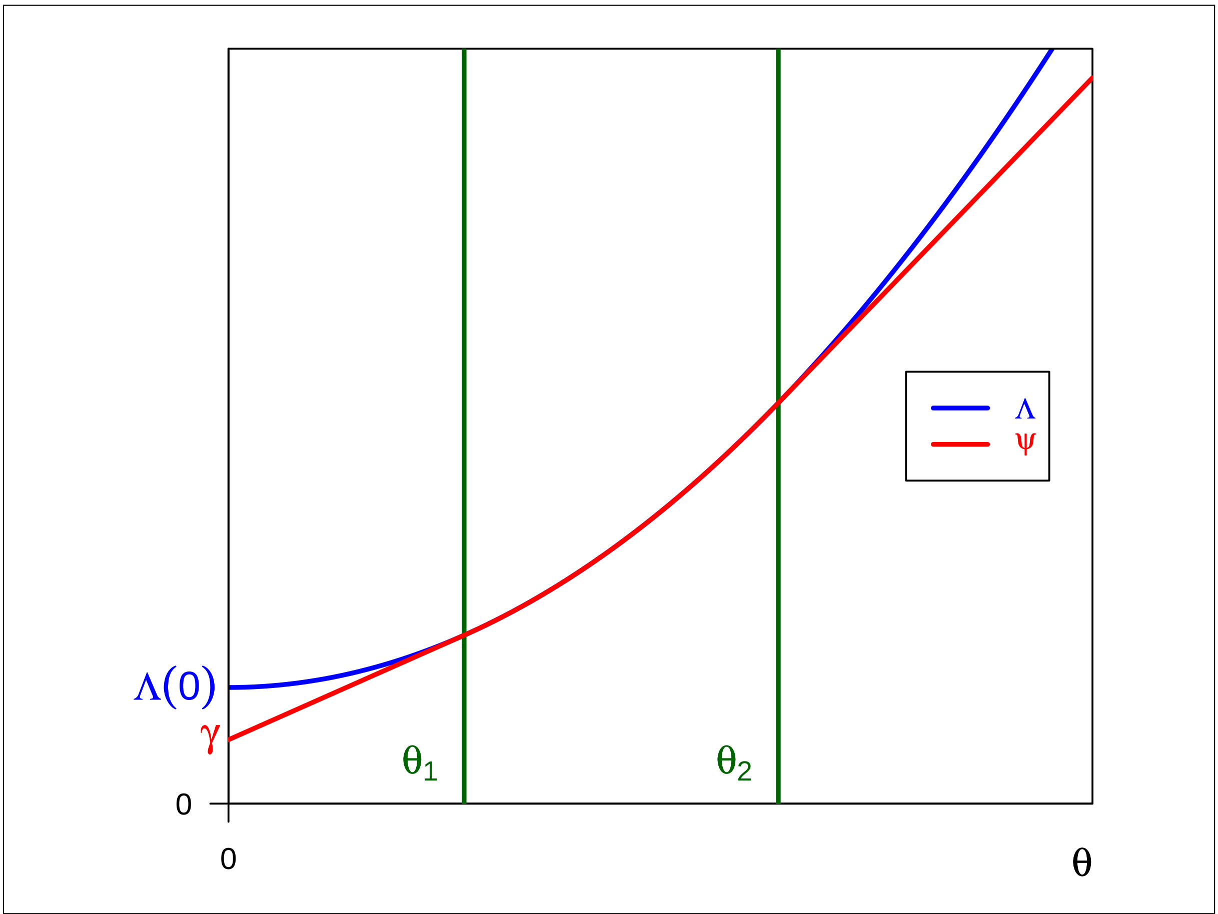


Figure 4.5: ψ and Λ in the case $\gamma < \Lambda(0)$ with $\theta_1 = \theta^*$ and $\theta_2 = \theta^\dagger$.

More explicitly, let Γ be defined as in Section 4.2. In Theorem 4.2.2, it is proved that Γ is the rate function of Z in the case of offspring distributions with geometrically bounded tails. Recall that another representation of Γ is

$$\Gamma(\theta) \quad = \quad \begin{cases} \gamma\left(1 - \frac{\theta}{\theta^*}\right) + \frac{\theta}{\theta^*}\Lambda(\theta^*) & , \quad \text{if } \theta < \theta^* \\ \Lambda(\theta) & , \quad \text{else} \end{cases}$$

where $0 \le \theta^* \le \infty$ is such that (see Definition (4.8))

$$\frac{\Lambda(\theta^*) - \gamma}{\theta^*} \quad = \quad \inf_{\theta \ge 0} \frac{\Lambda(\theta) - \gamma}{\theta} \, .$$

Let us define $\theta^\dagger$ by

$$\theta^\dagger \quad = \quad \sup\left\{\theta \ge \max\{0, \mathbb{E}[X]\} : \; \Gamma'(\theta) \le \beta \text{ and } \Gamma(\theta) < \infty\right\} . \tag{4.18}$$

Then

$$\psi(\theta) \quad = \quad \begin{cases} \Gamma(\theta) & , \quad \text{if } \theta \le \theta^\dagger \\ \beta\theta - \log(\mathbb{E}[e^{\beta X}]) & , \quad \text{else} \end{cases} \quad . \tag{4.19}$$

Phase Transitions

We will now briefly describe the different strategies associated with the corresponding phase transitions (of order two) of the rate function. Assume $0 < \theta^* < \theta^\dagger$ (see Figure 4.6 for an illustration):

- For $\theta < \theta^*$, the rate function ψ is identical to Γ. This means that no jumps occur. The best strategy is the same as in the case of offspring distributions without heavy tails (see Section 4.2). Namely, the process just survives until a time $\lfloor t_\theta n \rfloor$, $t_\theta \in (0,1)$ and then follows its expectation within good a environment (i.e. the associated random walk grows linearly such that $S_n - S_{\lfloor t_\theta n \rfloor} \approx \theta n$). Also recall that the slope of the linear growth of the associated random walk is always $\theta^* = \theta/(1-t_\theta)$.

- For $\theta^* \le \theta \le \theta^\dagger$, ψ is identical to Λ. Thus the best strategy is geometric growth (corresponding to linear growth of the associated random walk, such that $S_n \approx \theta n$) and the process (conditioned on the environment) just follows its expectation.

- For $\theta > \theta^\dagger$, ψ does not depend on the cost of survival γ and there is no initial period when the process just survives. The optimald strategy here is to jump at the beginning: $Z_1 \approx \exp(s_\theta n)$ and then follow the expectation in a good environment (i.e. $S_n \approx (\theta - s_\theta)n$). The slope of the linear growth of the associated random walk is always $\theta^\dagger = \theta - s_\theta$.

  To show this, we assume for convenience that $\gamma = \Lambda(0)$ (i.e. $\psi(\theta) = \inf_{s\in[0,\theta]}\{\beta s + \Lambda(\theta - s)\}$) and differentiability of Λ. Then the infimum in (4.16) is taken in s_θ such that

  $$\beta - \Lambda'(\theta - s_\theta) = 0 \ .$$

  As $\Lambda'(\theta^\dagger) = \beta$, the infimum is taken for $s_\theta = \theta - \theta^\dagger$ and the linear slope of the associated random walk is $\theta^\dagger$. Thus, for larger θ, larger values of Z_n are not realised by a stronger growth of the associated random walk but by a larger jump at the beginning.

The case $\theta^\dagger = 0$ corresponds to $\psi(\theta) = \gamma + \beta\theta$. Here the optimal strategy consists in just surviving until the end. In one of the last generations, an individual has $e^{\theta n}$-many offspring.

Note that in the case of $0 < \theta = \theta^* = \theta^\dagger$, the best strategy is no longer 'unique'. For any $t \in (0,1]$, there exists a $s \in [0,\theta]$ such that the optimal strategy is to just survive until time $\lfloor tn \rfloor$, then jump to $\exp(sn)$ and then follow a linear growth of the environment such that $S_n - S_{\lfloor tn \rfloor} \approx (\theta - s)n$ (see Figure 4.4).

4.3.2 Proof of the lower bound of Theorem 4.3.1

For the proof of the lower bound of Theorem 4.3.1, the following result is needed. It ensures that exceptional growth of the population can at least be achieved thanks to some suitable good environment sequence, whose probability decreases exponentially following the rate function of the random walk $(S_n)_{n\in\mathbb{N}}$. This result generalizes Proposition 1 in [BB09] for an exponential initial number of individuals. With a slight abuse of notation, we write below for the initial number of individuals $\exp(sn)$ instead of the integer part $\lfloor \exp(sn) \rfloor$.

Proposition 4.3.3. *Under Assumption $\mathcal{H}(\beta)$, for all $\theta \ge 0$ and $0 \le s \le \theta$,*

$$\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}_{\exp(sn)}(Z_n \ge e^{\theta n}) \quad \ge \quad -\Lambda\big((\theta - s)+\big) \ .$$

Proof. Recall that the rate function of the associated random walk is, for every $\theta' \ge 0$, defined by

$$\Lambda(\theta') = \sup_{\lambda \ge 0} \left\{ \lambda\theta' - \log \mathbb{E}[e^{\lambda X}] \right\} \ .$$

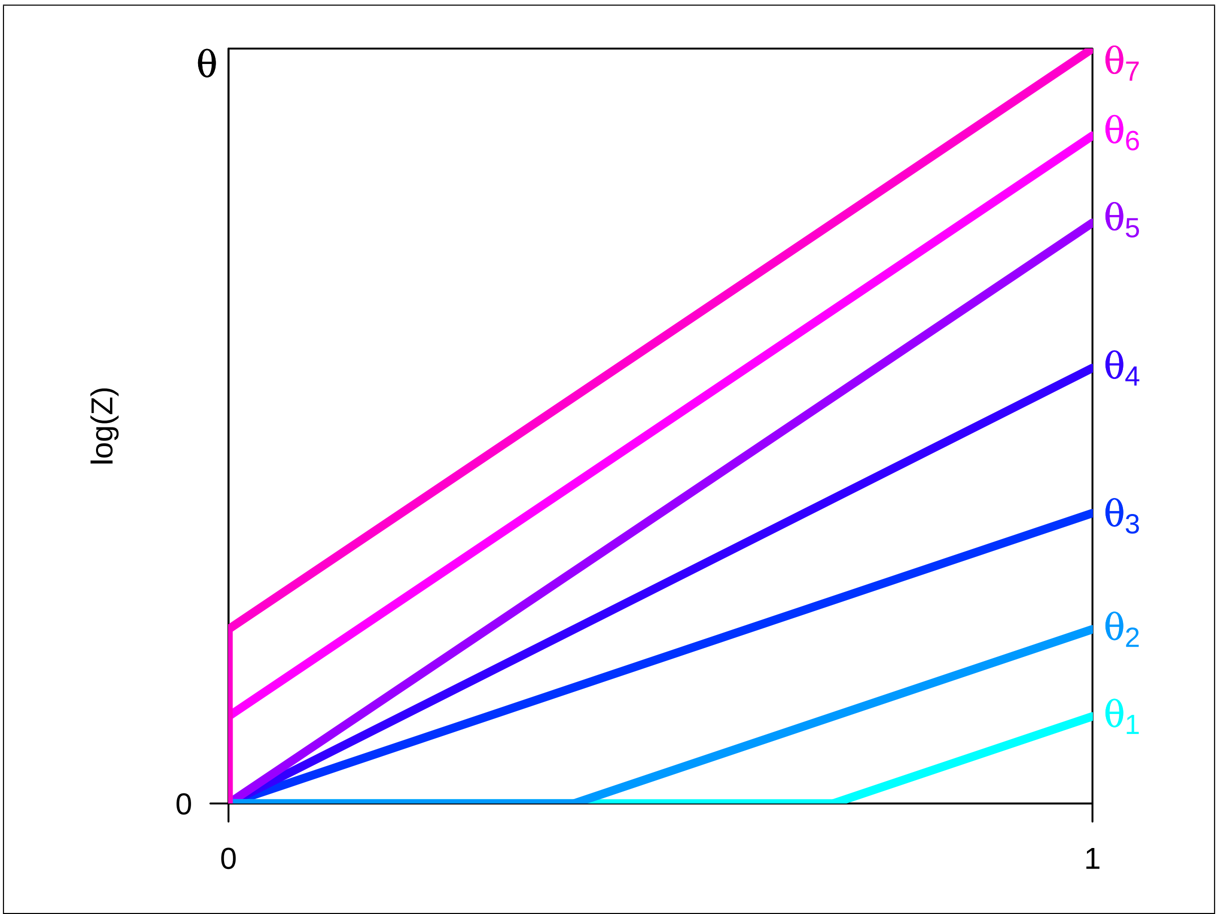


Figure 4.6: Most probable paths in the strongly subcritical case ($\gamma < \Lambda(0)$) for different θ with $\theta_1 < \theta_2 < \theta_3 = \theta^* < \theta_4 < \theta_5 = \theta^\dagger < \theta_6 < \theta_7$.

First, we assume that $\mathbb{E}[e^{\lambda X}] < \infty$ for every $\lambda \geq 0$. Then the derivative of $\lambda \to \mathbb{E}[e^{\lambda X}]$ exists for every $\lambda \geq 0$ and the supremum above is taken in $\lambda = \lambda_{\theta'}$ such that

$$\theta' = \frac{\mathbb{E}[Xe^{\lambda_{\theta'} X}]}{\mathbb{E}[e^{\lambda_{\theta'} X}]} \ .$$

Such as in the proof of Theorem 4.2.2, a change of measure is used. Introduce the probability measure $\widetilde{\mathbb{P}}$ defined by

$$\widetilde{\mathbb{P}}(X \in \mathrm{d}x) = \frac{e^{\lambda_{\theta'} x}}{\mathbb{E}[e^{\lambda_{\theta'} X}]} \ \mathbb{P}(X \in \mathrm{d}x) \ .$$

Under this new measure, $(S_n)_{n\in\mathbb{N}}$ is a random walk with drift $\widetilde{\mathbb{E}}[X] = \theta' > 0$ and Z_n is a supercritical BPRE.

For all $n \geq 1$, $\theta \in [0, \theta')$ and $\epsilon > 0$,

$$\begin{aligned}
&\mathbb{P}_{\exp(sn)}\big(Z_n \geq e^{(\theta+s)n}\big) \\
&\quad \geq \mathbb{P}_{\exp(sn)}\big(Z_n \geq e^{(\theta+s)n};\ S_n \leq (\theta'+\epsilon)n\big) \\
&\quad = \mathbb{E}[\exp(\lambda_{\theta'} X)]^n \widetilde{\mathbb{E}}_{\exp(sn)}\big[e^{-\lambda_{\theta'} S_n} \mathbb{1}_{\{S_n \leq [\theta'+\epsilon]n,\ Z_n \geq e^{(\theta+s)n}\}}\big] \\
&\quad \geq \exp\big(n\big[\log(\mathbb{E}[e^{\lambda_{\theta'} X}]) - \lambda_{\theta'}(\theta'+\epsilon)\big]\big)\widetilde{\mathbb{P}}_{\exp(sn)}\big(Z_n \geq e^{(\theta+s)n},\ S_n \leq (\theta'+\epsilon)n\big) \\
&\quad \geq \exp(n(-\Lambda(\theta') - \lambda_{\theta'}\epsilon))\big[\widetilde{\mathbb{P}}_{\exp(sn)}\big(Z_n \geq e^{(\theta+s)n}\big) - \widetilde{\mathbb{P}}\big(S_n > (\theta'+\epsilon)n\big)\big].
\end{aligned}$$

As $\widetilde{\mathbb{P}}\big(S_n > (\theta' + \epsilon)n\big) \to 0$ when $n \to \infty$, it suffices to prove that

$$\liminf_{n\to\infty} \widetilde{\mathbb{P}}_{\exp(sn)}\big(Z_n \geq e^{(\theta+s)n}\big) > 0, \tag{4.20}$$

and Proposition 4.3.3 follows by letting $\epsilon \to 0, \theta' \to \theta$.

Relation (4.20) results from the fact that under $\widetilde{\mathbb{P}}$ the population Z_n starting from one single individual grows as $e^{S_n} \asymp n\theta'$ on the nonextinction event. More precisely, individuals of the initial population are labeled and the number of descendants in generation n of individual i is denoted by $Z_n^{(i)}$. For given $N \in \mathbb{N}$, introduce the 'success' probability p_n:

$$p_n = \mathbb{P}_1(Z_n \geq N e^{n\theta} \mid \Pi) \quad \text{a.s.}$$

Then, conditioned on Π, and for $N \geq 1$, the number of initial individuals whose number of descendants in generation n is larger than $Ne^{n\theta}$,

$$N_n := \#\{1 \leq i \leq e^{sn} : Z_n^{(i)} \geq N e^{n\theta}\} \quad \text{a.s. },$$

follows a binomial distribution with parameters (e^{sn}, p_n). Moreover, as $\mathbb{E}[N_n \mid \Pi] = e^{sn}p_n$ a.s.,

$$\widetilde{\mathbb{P}}_{\exp(sn)}\big(Z_n \geq e^{(\theta+s)n}\big) \geq \widetilde{\mathbb{P}}_{\exp(sn)}\big(N_n \geq e^{sn}/N\big) \geq \widetilde{\mathbb{P}}_{\exp(sn)}\Big(N_n \geq \frac{\mathbb{E}[N_n \mid \Pi]}{Np_n}\Big) .$$

Applying the classical inequality due to Paley and Zygmund again for $r \in [0, 1]$ (compare proof of Lemma 4.2.7 on page 52 or e.g. [Kal01, p. 63]) yields

$$\mathbb{P}(Y \geq r\mathbb{E}[Y]) \quad \geq \quad (1-r)^2 \frac{\mathbb{E}[Y]^2}{\mathbb{E}[Y^2]} . \tag{4.21}$$

Adding that $\mathbb{E}[N_n \mid \Pi] = e^{sn}p_n$ and $\mathbb{E}[N_n^2 \mid \Pi] = e^{2sn}p_n^2 + e^{sn}p_n(1 - p_n)$ a.s., we get

$$\widetilde{\mathbb{P}}_{\exp(sn)}\Big(N_n \geq \frac{\mathbb{E}[N_n \mid \Pi]}{Np_n} \,\Big|\, \Pi\Big) \geq \Big[1 - 1 \wedge \frac{1}{Np_n}\Big]^2 \frac{\mathbb{E}[N_n \mid \Pi]^2}{\mathbb{E}[N_n^2 \mid \Pi]} \geq \frac{\Big[1 - 1 \wedge \frac{1}{Np_n}\Big]^2}{1 + \frac{e^{-sn}}{p_n}} \quad \text{a.s.}$$

U
nder Assumption $\mathcal{H}(\beta)$,

$$\widetilde{\mathbb{E}}\Big[\sum_{k\in\mathbb{N}} k^{1+\epsilon}\mathcal{P}(R = k)/m\Big] \leq \widetilde{\mathbb{E}}\Big[\sum_{k\in\mathbb{N}} k^{1+\epsilon}\mathcal{P}(R = k | R > 0)\Big] < \infty ,$$

for every $0 < \epsilon < \beta - 1$. Theorem 3 in [GL01] ensures that for every $N \in \mathbb{N}$,

$$\widetilde{\mathbb{E}}[p_n] = \widetilde{\mathbb{P}}_1(Z_n \geq N e^{\theta n}) \overset{n\to\infty}{\longrightarrow} \widetilde{\mathbb{P}}_1(\forall n \in \mathbb{N} : Z_n > 0) > 0 .$$

As the right-hand side does not depend on $N \geq 1$ and $p_n \leq 1$, for N large enough,

$$\delta := \liminf_{n\to\infty} \widetilde{\mathbb{P}}(p_n \geq 2/N) > 0$$

and thus

$$\liminf_{n\to\infty} \widetilde{\mathbb{P}}_{\exp(sn)}\big(Z_n \geq e^{(\theta+s)n}\big) \geq \liminf_{n\to\infty} \widetilde{\mathbb{E}}\Big[\frac{\big[1 - 1 \wedge 1/Np_n\big]^2}{1 + 1/p_n}\Big] \geq \frac{\delta(1 - 1/2)^2}{1 + N/2} > 0 .$$

This proves (4.20) and ends up the proof when $\mathbb{E}[e^{\lambda X}] < \infty$ for every $\lambda \geq 0$. The general case follows by a standard approximation argument (the same arguments as in the proof of Lemma 4.2.7, page 53, can be applied). □

Proof of the lower bound in Theorem 4.3.1. The proof now amounts to exhibiting good trajectories which realize the large deviation event $\{Z_n \geq e^{\theta n}\}$. For every $t \in (0, 1)$ and $s \in [0, \theta]$, by Markov property,

$$\mathbb{P}(Z_n \geq e^{\theta n}) \geq \mathbb{P}(Z_{\lfloor tn \rfloor} > 0)\mathbb{P}(Z_1 \geq e^{sn})\mathbb{P}_{\exp(sn)}(Z_{n-\lfloor tn \rfloor} \geq e^{\theta n}) .$$

Taking the logarithm, by (4.1) the first term can be written as

$$\frac{1}{tn}\log(\mathbb{P}(Z_{\lfloor tn\rfloor}>0)) \overset{n\to\infty}{\longrightarrow} -\gamma\ .$$

Using that $\log(\mathbb{P}(Z_1>z))/\log(z) \overset{z\to\infty}{\longrightarrow} -\beta$ yields for the second term

$$\frac{1}{n}\log(\mathbb{P}(Z_1\geq e^{sn})) \overset{n\to\infty}{\longrightarrow} -s\beta\ .$$

Finally, by Proposition 4.3.3,

$$\liminf_{n\to\infty}\frac{1}{(1-t)n}\log\left(\mathbb{P}_{\exp(sn)}(Z_{n-[tn]}\geq e^{\theta n})\right)\geq -\Lambda((\theta-s)/(1-t)+)$$

since

$$\mathbb{P}_{\exp(sn)}(Z_{n-[tn]}\geq e^{\theta n}) \;=\; \mathbb{P}_{\exp((1-t)ns/(1-t))}(Z_{n-[tn]}\geq e^{n(1-t)\theta/(1-t)})\ .$$

Combining the first inequality and the last three limits ensures that

$$\liminf_{n\to\infty}\frac{1}{n}\log(\mathbb{P}(Z_n\geq e^{\theta n}))\geq -\inf_{t\in[0,1],s\in[0,\theta]}\Big\{t\gamma+\beta s+(1-t)\Lambda((\theta-s)/(1-t)+)\Big\}\ .$$

As a convex nonnegative function, Λ has at most one jump (to infinity). Thus the infimum above is $\psi(\theta)$. To see this, we only have to consider the jump point. Say, there are $s_\theta\in[0,\theta]$ and $t_\theta\in[0,1)$ such that

$$t_\theta\gamma+\beta s_\theta+(1-t_\theta)\Lambda((\theta-s_\theta)/(1-t_\theta)) \;=\; \psi(\theta)<\infty$$

and $\Lambda((\theta-s_\theta)/(1-t_\theta)+)=\infty$. Then, as $(\theta-s_\theta)/(1-t_\theta)$ is the only jump point, for any $\epsilon>0$ there is a $\delta>0$ such that

$$\begin{aligned}\psi(\theta)-\epsilon \;&\leq\; t_\theta\gamma+\beta(s_\theta-\delta)+(1-t_\theta)\Lambda((\theta-s_\theta-\delta)/(1-t_\theta)+)\\ &=\; t_\theta\gamma+\beta(s_\theta-\delta)+(1-t_\theta)\Lambda((\theta-s_\theta-\delta)/(1-t_\theta)).\end{aligned}$$

Finally letting $\epsilon\to 0$ proves the result and thereby the lower bound of Theorem 4.3.1. □

4.3.3 Proof of the upper bound of Theorem 4.3.1 for $\beta\in(1,2]$

Recall the definition of the minimum (resp. maximum) of the associated random walk[21] up to time n,

$$L_n \;:=\; \min_{0\leq k\leq n} S_k \qquad , \qquad M_n := \max_{0\leq k\leq n} S_k$$

Using the inequality (compare introduction (1.3) or e.g. [BGK05])

$$\mathbb{P}(Z_n>0|\Pi)\leq e^{L_n}. \tag{4.22}$$

As explained in Chapter 1, under mild assumptions, the above estimate yields the correct exponential decay rate:

$$\gamma \;=\; -\lim_{n\to\infty}\frac{1}{n}\log\mathbb{P}(Z_n>0) = -\lim_{n\to\infty}\frac{1}{n}\log\mathbb{E}\big[e^{L_n}\big].$$

In Lemma 4.3.5, the above relation is generalized and proved rigorously under Assumption $\mathcal{H}(\beta)$.

The following theorem, which is an equivalent of Theorem 4.2.4 for heavy tailed offspring distributions, provides the key bound for the tail probabilities of Z_n, conditioned on the environment.

Theorem 4.3.4. *Under Assumption $\mathcal{H}(\beta)$ for some $\beta\in(1,2]$, there exist a constant $0<c<\infty$ and a positive nondecreasing and slowly varying function Υ such that for all $k\geq 1$ and $n\geq 1$,*

$$\mathbb{P}(Z_n>k|\Pi) \;\leq\; c\; n\cdot\Upsilon\big(n^{2/(\beta-1)}e^{M_n-L_n}k\big)\cdot e^{L_n}\cdot(e^{S_n-L_n}/k)^\beta \quad a.s.$$

[21] Here, we consider the infimum over $0\leq k\leq n$ instead of $1\leq k\leq n$ as in the definition in Chapter 1.

Let us briefly explain this result. The probability to survive until time n evolves as e^{L_n}. Conditioned on survival, a good environment corresponds to large values of $(S_n - L_n)$. The possibility of high reproduction of the initial individual is reflected by the last term, $k^{-\beta}$. Conditioned on the environment and survival, the expected size of the process at time n is of order $e^{S_n-L_n}$: this corresponds to the 'best time period' for the growth of the process. Thus, this theorem essentially says that conditioned on $Z_n > 0$, the tail distribution of $Z_n/e^{S_n-L_n}$ is at most polynomial with exponent $-\beta$.

Recall the notation from Section 4.2.1, i.e. that $f_n(s) = \sum_{k=0}^{\infty} s^k Q_n(k)$ is the probability generating function of the offspring distribution of an individual in generation $n-1$, and (see (3.11))

$$\mathbb{E}[s^{Z_n}|\Pi] \;=\; f_1(f_2(\cdots f_n(s)\cdots)) \;=\; f_{0,n}(s) \quad \text{a.s.} \quad (0 \le s \le 1)\ .$$

For the proofs, it is suitable to work with an alternative expression, namely for every $k \ge 1$,

$$g_k(s) := \frac{1-f_k(s)}{1-s} \quad \text{a.s.} \quad (0 \le s \le 1)$$

and

$$g_{0,n}(s) \;:=\; \sum_{k=0}^{\infty} s^k \mathbb{P}(Z_n > k|\Pi) \;=\; \frac{1-f_{0,n}(s)}{1-s} \quad \text{a.s.} \quad (0 \le s \le 1)\ . \tag{4.23}$$

Moreover, we need the following auxiliary function defined for every $\mu \in (0,1]$ by

$$h_{\mu,k}(s) \;:=\; \frac{1}{(1-f_k(s))^\mu} - \frac{1}{(f_k'(1)(1-s))^\mu} = \frac{g_k(1)^\mu - g_k(s)^\mu}{(g_k(1)g_k(s)(1-s))^\mu} \quad \text{a.s.} \quad (0 \le s \le 1)\ . \tag{4.24}$$

Finally, we define for all $1 \le k \le n$,

$$\begin{aligned} U_k \;&:=\; (f_1'(1)\cdots f_k'(1))^{-1} \;=\; f_{0,k}'(1)^{-1} \;=\; e^{-S_k}, \\ f_{k,n} \;&:=\; f_{k+1}\circ f_{k+2}\circ\cdots\circ f_n, \quad 0 \le k < n;\ f_{n,n} = id \quad \text{a.s.} \end{aligned}$$

By a telescope summation argument, we derive a formula similar to (3.13).

$$\begin{aligned} \frac{1}{(1-f_{0,n}(s))^\mu} \;&=\; \frac{U_0^\mu}{(1-f_{0,n}(s))^\mu} \\ &=\; \frac{U_n^\mu}{(1-f_{n,n}(s))^\mu} + \sum_{k=0}^{n-1}\left(\frac{U_k^\mu}{(1-f_{k,n}(s))^\mu} - \frac{U_{k+1}^\mu}{(1-f_{k+1,n}(s))^\mu}\right) \\ &=\; \frac{U_n^\mu}{(1-s)^\mu} + \sum_{k=0}^{n-1} U_k^\mu \left(\frac{1}{(1-f_{k+1}(f_{k+1,n}(s)))^\mu} - \frac{1}{(f_{k+1}'(1)(1-f_{k+1,n}(s)))^\mu}\right) \\ &=\; \frac{U_n^\mu}{(1-s)^\mu} + \sum_{k=0}^{n-1} U_k^\mu h_{\mu,k+1}(f_{k+1,n}(s)), \quad s \ge 0\ . \end{aligned} \tag{4.25}$$

Proof of Theorem 4.3.4. In the same vein as Theorem 4.2.4, an upper bound for $\mathbb{P}(Z_n > z|Pi)$ is obtained from the divergence of

$$g_{0,n}'(s) = \sum_{j=0}^{\infty} j\mathbb{P}(Z_n > j|\Pi)s^{j-1}$$

as $s \to 1$. In that purpose, by (4.25) for $\mu = \beta - 1$,

$$g_{0,n}(s) \;=\; \left(U_n^{\beta-1} + (1-s)^{\beta-1}\sum_{k=0}^{n-1} U_k^{\beta-1} h_{\beta-1,k+1}(f_{k+1,n}(s))\right)^{-1/(\beta-1)} \quad (0 \le s \le 1) \quad \text{a.s.}$$

The first derivative of $g_{0,n}$ is calculated as follows:

$$
\begin{aligned}
& g'_{0,n}(s) \\
= & -(\beta-1)^{-1}\Big(U_n^{\beta-1}+(1-s)^{\beta-1}\sum_{k=0}^{n-1}U_k^{\beta-1}h_{\beta-1,k+1}(f_{k+1,n}(s))\Big)^{-1-1/(\beta-1)} \\
& \times\Big(-(\beta-1)(1-s)^{\beta-2}\sum_{k=0}^{n-1}U_k^{\beta-1}h_{\beta-1,k+1}(f_{k+1,n}(s)) \\
& \qquad +(1-s)^{\beta-1}\sum_{k=0}^{n-1}U_k^{\beta-1}h'_{\beta-1,k+1}(f_{k+1,n}(s))f'_{k+1,n}(s)\Big) \\
= & \frac{\sum_{k=0}^{n-1}U_k^{\beta-1}h_{\beta-1,k+1}(f_{k+1,n}(s))-(\beta-1)^{-1}(1-s)\sum_{k=0}^{n-1}U_k^{\beta-1}h'_{\beta-1,k+1}(f_{k+1,n}(s))f'_{k+1,n}(s)}{(1-s)^{2-\beta}\Big(U_n^{\beta-1}+(1-s)^{\beta-1}\sum_{k=0}^{n-1}U_k^{\beta-1}h_{\beta-1,k+1}(f_{k+1,n}(s))\Big)^{1+1/(\beta-1)}} \\
\leq & \frac{\sum_{k=0}^{n-1}U_k^{\beta-1}\Big(h_{\beta-1,k+1}(f_{k+1,n}(s))-(\beta-1)^{-1}h'_{\beta-1,k+1}(f_{k+1,n}(s))f'_{k+1,n}(s)(1-s)\Big)}{U_n^{\beta}(1-s)^{2-\beta}} \;. \qquad (4.26)
\end{aligned}
$$

Now Lemma 4.3.9 in Section 4.3.7 ensures that there exists a $c>0$ such that for every $s\in[0,1)$,

$$
\begin{aligned}
h_{\beta-1,k}(s) &\leq c\Upsilon(1/(1-s)) \;, \\
-h'_{\beta-1,k}(s) &\leq c\Upsilon(1/(1-s))/(1-s) \quad \text{a.s.} \qquad (4.27)
\end{aligned}
$$

Moreover, using (4.25), Lemma 4.3.9 in Section 4.3.7 for $0<\mu<\beta-1$ and $U_k\leq e^{-L_n}$ for every $0\leq k\leq n$, there exists a $c\geq 1$ such that

$$
\frac{1}{(1-f_{k+1,n}(s))^{\mu}} \leq \frac{e^{-\mu(S_n-S_k)}}{(1-s)^{\mu}}+n\;c\;e^{-\mu(S_n-S_k)}\leq ce^{\mu(M_n-L_n)}(n+1)/(1-s)^{\mu} \quad (0\leq s\leq 1)\;.
$$

Combining this inequality with (4.27) yields

$$
h_{\beta-1,k+1}(f_{k+1,n}(s)) \leq c\Upsilon\big((n+1)^{1/\mu}e^{M_n-L_n}(1-s)^{-1}\big) \quad (0\leq s<1) \quad \text{a.s.}
$$

for some $c>0$. By convexity of $f_{k+1,n}$, $f_{k+1,n}(s)\leq 1-f'_{k+1,n}(s)(1-s)$ and thus (4.27) ensures that

$$
\begin{aligned}
-h'_{\beta-1,k+1}(f_{k+1,n}(s))f'_{k+1,n}(s)(1-s) &\leq c\;f'_{k+1,n}(s)(1-s)\Upsilon\big(1/(1-f_{k+1,n}(s))\big)\frac{1}{1-f_{k+1,n}(s)} \\
&\leq c\;\Upsilon\big((n+1)^{1/\mu}e^{M_n-L_n}/(1-s)\big) \quad (0\leq s<1) \quad \text{a.s.}
\end{aligned}
$$

Using the two last estimates with $\mu=(\beta-1)/2$ together in (4.26) yields

$$
g'_{0,n}(s) \leq \frac{c\;n\;e^{-(\beta-1)L_n}\Upsilon\big((n+1)^{2/(\beta-1)}e^{M_n-L_n}(1-s)^{-1}\big)}{U_n^{\beta}(1-s)^{2-\beta}} \quad (0\leq s\leq 1) \quad \text{a.s.}
$$

From below, we estimate $g'_{0,n}(s)$ for all $k\geq 1$ and $s\in[0,1]$ by

$$
\begin{aligned}
g'_{0,n}(s) &\geq \sum_{j=1}^{k} j\mathbb{P}(Z_n>j|\Pi)s^{j-1} \\
&\geq s^k\frac{k^2}{2}\mathbb{P}(Z_n>k|\Pi)\;. \qquad (4.28)
\end{aligned}
$$

By setting $s=1-1/k$ in the two last inequalities, we get

$$
\Big(1-\frac{1}{k}\Big)^k\frac{k^2}{2}\mathbb{P}(Z_n>k|\Pi) \leq \frac{c\;n\;e^{-(\beta-1)L_n}k^{2-\beta}\;\Upsilon\big(k(n+1)^{2/(\beta-1)}e^{M_n-L_n}\big)}{U_n^{\beta}}\;,
$$

which ends up the proof since $U_n=e^{-L_n}$. $\square$

For the proof of the upper bound in Theorem 4.3.1, the following characterization of the survival cost is required:

Lemma 4.3.5. *Under Assumption $\mathcal{H}(\beta)$, for all $\theta \geq 0, b \geq 0$ and Υ positive nondecreasing and slowly varying at infinity,*

$$\gamma \quad = \quad -\lim_{n\to\infty} \frac{1}{n} \log \mathbb{E}\big[\Upsilon(n^b e^{\theta n} e^{-L_n}) e^{L_n}\big] \ .$$

Proof of Lemma 4.3.5. First let $\Upsilon = 1$. We use (4.25) with some $0 < \mu < \beta - 1$ and by Lemma 4.3.9 proved in Section 4.3.7

$$\begin{aligned} \mathbb{P}(Z_n > 0|\Pi) \quad &\geq \quad \frac{1}{(e^{-\mu S_n} + \sum_{k=0}^{n-1} e^{-\mu S_k} h_{\mu,k+1}(f_{k+1,n}(1)))^{\mu^{-1}}} \\ &\geq \quad n^{-1/\mu}\ c^{-1} e^{L_n}. \end{aligned} \tag{4.29}$$

The upper bound is already proved by (4.22) and we get

$$\gamma \quad = \quad -\lim_{n\to\infty} \frac{1}{n} \log \mathbb{E}\big[e^{L_n}\big] \ .$$

As Υ is nondecreasing,

$$\gamma \quad \geq \quad \limsup_{n\to\infty} -\frac{1}{n} \log \mathbb{E}[\Upsilon(n^b e^{\theta n} e^{-L_n}) e^{L_n}] \ .$$

For the converse inequality, we use that $\mathbb{E}\big[e^{tL_n}\big]$ is nonincreasing in n to define

$$\chi(t) \quad := \quad -\lim_{n\to\infty} \frac{1}{n} \log \mathbb{E}\big[e^{tL_n}\big] \ .$$

Note that $\chi(t) \geq 0$ and by [dH00, Lemma V.4], $\chi(t)$ is finite and convex. Thus χ is continuous.

By properties of slowly varying sequences (see appendix or e.g. [BGT87, Proposition 1.3.6, p. 16]), for any $\delta > 0$, $x^{-\delta}\Upsilon(x) \to 0$ as $x \to \infty$ and

$$-\lim_{n\to\infty} \frac{1}{n} \log \mathbb{E}\big[\Upsilon(n^b e^{\theta n} e^{M_n - L_n}) e^{L_n}\big] \geq -\delta\theta - \lim_{n\to\infty} \frac{1}{n} \log \mathbb{E}\big[e^{(1+\delta)L_n}\big].$$

Letting $\delta \to 0$ and using continuity of χ ends up the proof. □

Proof of the upper bound in Theorem 4.3.1. First, recall the following classical large deviation inequality (see e.g. [dH00]):

$$\mathbb{P}(S_n \geq \theta n) \ \leq \ e^{-\Lambda(\theta)n} \ . \tag{4.30}$$

Define the first time τ_n when the random walk $(S_i)_{i\leq n}$ reaches its minimum value on $[0,n]$ by (see (1.4))

$$\tau_n := \inf\{0 \leq k \leq n : \ S_k = L_n\} \ .$$

The probability of having an extraordinarily large population is decomposed according to $S_n - L_n$. o control the term in the slowly varying function in Theorem 4.3.4, we also add a term bounding the maximum of the random walk up to time n. Let $r \in \mathbb{N}$. Then

$$\begin{aligned} \mathbb{P}(Z_n \geq e^{\theta n}) \quad = \quad & \mathbb{P}(Z_n \geq e^{\theta n}, S_n - L_n \geq \theta n) + \mathbb{E}[\mathbb{P}(Z_n \geq e^{\theta n}|\Pi); S_n - L_n < \theta n, M_n \leq r\theta n] \\ & + \mathbb{E}[\mathbb{P}(Z_n \geq e^{\theta n}|\Pi); S_n - L_n < \theta n, M_n > r\theta n] \ . \end{aligned} \tag{4.31}$$

The asymptotic of the first term has already been studied in Section 4.2 on page 51. Recall that, using (4.30)

$$\begin{aligned} \mathbb{P}(Z_n \geq e^{\theta n},\ S_n - L_n \geq \theta n) \quad &\leq \quad \sum_{i=0}^{n} \mathbb{P}(Z_i > 0)\mathbb{P}(S_n - S_i \geq \theta n) \\ &\leq \quad \sum_{i=0}^{n} \mathbb{P}(Z_i > 0)\ e^{-(n-i)\Lambda(\theta n/(n-i))} \ . \end{aligned} \tag{4.32}$$

This ensures that

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}(Z_n \geq e^{\theta n}, S_n - L_n \geq \theta n) \quad \leq \quad -\Gamma(\theta) \quad , \tag{4.33}$$

where

$$\Gamma(\theta) \quad = \quad \inf_{0<t\leq 1} \{t\gamma + (1-t)\Lambda(\theta/(1-t))\} \ .$$

For the second term in (4.31), Theorem 4.3.4 is used and by Markov property of $(S_n)_{n\geq 0}$:

$$\begin{aligned}
&\mathbb{E}[\mathbb{P}(Z_n \geq e^{\theta n}|\Pi); S_n - L_n < \theta n, M_n \leq r\theta n] \\
&\quad \leq c\ n\ \mathbb{E}\Big[\Upsilon(n^{2/(\beta-1)} e^{M_n - L_n} e^{\theta n}) e^{L_n} e^{\beta(S_n - L_n - \theta n)}; S_n - L_n < \theta n, M_n \leq r\theta n\Big] \\
&\quad \leq c\ n\ \sum_{k=0}^{n} \mathbb{E}\Big[\Upsilon(n^{2/(\beta-1)} e^{-S_k} e^{(r+1)\theta n}) e^{S_k} e^{\beta(S_n - S_k - \theta n)}; S_n - L_n < \theta n, \tau_n = k\Big] \\
&\quad \leq c\ n\ \sum_{k=0}^{n} \mathbb{E}\Big[\Upsilon(n^{2/(\beta-1)} e^{-S_k} e^{(r+1)\theta n}) e^{S_k}; \tau_k = k\Big] \mathbb{E}\Big[e^{-\beta(\theta n - S_{n-k})}; S_{n-k} < \theta n, L_{n-k} \geq 0\Big]
\end{aligned}$$

Let $\epsilon = 1/n^2$ and $m_\epsilon = \lceil \theta/\epsilon \rceil$. Using that

$$\begin{aligned}
\mathbb{E}\big[\Upsilon(n^{2/(\beta-1)} e^{-S_k} e^{(r+1)\theta n}) e^{S_k}; \tau_k = k\big] \quad &= \quad \mathbb{E}\big[\Upsilon(n^{2/(\beta-1)} e^{-L_k} e^{(r+1)\theta n}) e^{L_k}, \tau_k = k\big] \\
&\leq \quad \mathbb{E}\big[\Upsilon(n^{2/(\beta-1)} e^{-L_k} e^{(r+1)\theta n}) e^{L_k}\big]
\end{aligned}$$

and it is deduced from(4.31),

$$\begin{aligned}
&\mathbb{E}[\mathbb{P}(Z_n \geq e^{\theta n}|\Pi); S_n - L_n < \theta n, M_n \leq r\theta n] \\
&\quad \leq \ c\ n\ \sum_{k=0}^{n} \mathbb{E}\big[\Upsilon(n^{2/(\beta-1)} e^{-L_k} e^{(r+1)\theta n}) e^{L_k}\big] \sum_{j=0}^{m_\epsilon} e^{-\beta(\theta-(j+1)\epsilon)n} \mathbb{P}\big(S_{n-k} \in [nj\epsilon, n(j+1)\epsilon), L_{n-k} \geq 0\big) \\
&\quad \leq \ c\ n\ \sum_{k=0}^{n} \mathbb{E}\big[\Upsilon(n^{2/(\beta-1)} e^{-L_k} e^{(r+1)\theta n}) e^{L_k}\big] \sum_{j=0}^{m_\epsilon} e^{-\beta(\theta-(j+1)\epsilon)n} e^{-\Lambda(j\epsilon n/(n-k))(n-k)} \\
&\quad \leq \ c\ \theta\ n^4 \sup_{0<t\leq 1, 0\leq s\leq \theta} \Big\{\mathbb{E}\Big[\Upsilon(n^{2/(\beta-1)} e^{-L_{\lfloor tn \rfloor}} e^{(r+1)\theta n}) e^{L_{\lfloor tn \rfloor}}\Big] \cdot e^{-(\beta s - (1-t)\Lambda((\theta-s)/(1-t)))n}\Big\} \ .
\end{aligned}$$

Together with Lemma 4.3.5, this yields for every $r \in \mathbb{N}$,

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{E}\big[\mathbb{P}(Z_n \geq e^{\theta n}|\Pi); S_n - L_n < \theta n, M_n \leq r\theta n\big] \quad \leq \quad -\psi(\theta), \tag{4.34}$$

where

$$\psi(\theta) \quad = \quad \inf_{0<t\leq 1, 0\leq s\leq \theta} \Big\{t\gamma + \beta s + (1-t)\Lambda((\theta-s)/(1-t))\Big\} \ .$$

As to the third term in (4.31), by duality,

$$\begin{aligned}
\mathbb{E}[\mathbb{P}(Z_n \geq e^{\theta n}|\Pi); S_n - L_n < \theta n, M_n > r\theta n] &\leq \mathbb{P}(M_n > r\theta n) \\
&= \mathbb{P}(\max_{k=0,\ldots,n} (S_n - S_k) > r\theta n) = \mathbb{P}(S_n - L_n > r\theta n).
\end{aligned} \tag{4.35}$$

It has been proved in Section 4.2 that,

$$\Gamma_0(x) := \limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}(S_n - L_n > xn) = \inf_{0<t\leq 1} \big\{(1-t)\Lambda(x/(1-t))\big\} \stackrel{x\to\infty}{\longrightarrow} \infty.$$

Combining this result with (4.31), (4.33), (4.34) and (4.35) shows that

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}(Z_n \geq e^{\theta n}) \ \leq \ -\min\{\Gamma(\theta); \psi(\theta); \Gamma_0(r\theta)\}.$$

Observe that $\psi(\theta) \leq \Gamma(\theta)$ since the infimum is considered on a larger set for ψ than for Γ. Adding that $\Gamma_0(x) \to \infty$ as $x \to \infty$, we get

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}(Z_n \geq e^{\theta n}) \quad \leq \quad -\psi(\theta) \ ,$$

by letting $r \to \infty$. This proves the upper bound of Theorem 4.3.1. □

4.3.4 Adaptation of the proof of the upper bound for $\beta > 2$

First, Lemma 4.3.5 still holds for $\beta > 2$ by following the same proof. Indeed, using (4.25) for $\mu = 1$ together with Lemma 4.3.9 ensures that

$$\mathbb{P}(Z_n > 0|\Pi) = 1 - f_{0,n}(0) \quad \geq \quad \frac{1}{e^{-S_n} + \sum_{k=0}^{n-1} e^{-S_k} h_{k+1}(f_{k+1,n}(0))} \geq n^{-1}\ c^{-1} e^{L_n}.$$

The main difficulty is to obtain an equivalent of Theorem 4.3.4. For this, the calculation of higher order derivatives of $g_{0,n}$ is needed and the upper bound on tail probabilities of Z_n contains an additional term:

Theorem 4.3.6. *Under Assumption $\mathcal{H}(\beta)$ for some $\beta > 2$, there are a constant $0 < c < \infty$ and a positive nondecreasing slowly varying function Υ such that for every $k \geq 1$,*

$$P(Z_n > k|\Pi) \ \leq \ c\ e^{S_n} n^{\beta} \Upsilon(n^2 e^{M_n - L_n} k) \max\left\{k^{-\beta} e^{(\beta-1)(S_n-L_n)}; k^{-\lceil\beta\rceil-1} e^{\lceil\beta\rceil(S_n-L_n)}\right\} \quad a.s.$$

In the following, the l^{th} derivative of a function f is denoted by $f^{(l)}$. For the proof, we use the functions

$$h_k(s) = \frac{1}{(1-f_k(s))} - \frac{1}{f_k'(1)(1-s)} = \frac{g_k(1) - g_k(s)}{g_k(1) g_k(s)(1-s)} \quad \text{a.s.} \quad (0 \leq s < 1)$$

and

$$H(s) = \sum_{k=0}^{n-1} U_k h_{k+1}(f_{k+1,n}(s)) \quad \text{a.s.} \quad (0 \leq s < 1). \tag{4.36}$$

Then (4.25) with $\mu = 1$ yields

$$g_{0,n}(s)^{-1} \quad = \quad \frac{1-s}{1-f_{0,n}(s)} \ = \ U_n + (1-s)H(s)$$

and calculating the l-th derivative of the above equation, we get for all $l \geq 1$ and $s \in [0,1)$,

$$\frac{d^l}{ds^l} g_{0,n}(s)^{-1} \quad = \quad (1-s)H^{(l)}(s) - lH^{(l-1)}(s) \quad \text{a.s.} \quad (0 \leq s \leq 1). \tag{4.37}$$

The proof of the upper bound is organized as follows. First, we prove the following technical lemma, which provides useful bounds for power generating series. Then Theorem 4.3.6 is derived. Finally the main lines of the proof of the upper bound of Theorem 4.3.1 for $\beta > 2$ are explained (following the proof for $\beta \in (1,2]$). For simplicity of notation, we introduce $\leq_c$ which means that the inequality is fulfilled up to a multiplicative constant c that does not depend on s, k or Π.

Lemma 4.3.7. *Under Assumption $\mathcal{H}(\beta)$, for every $l \leq \lceil\beta\rceil - 1$,*

$$f_{0,n}^{(l)}(1) \quad \leq_c \quad n^{l-1}\ e^{S_n} e^{(l-1)(S_n-L_n)} \quad a.s. \tag{4.38}$$

Moreover, the following estimates hold a.s. for every $s \in [0,1)$ respectively for $l < \lceil\beta\rceil - 2$, $l = \lceil\beta\rceil - 2$ and $l = \lceil\beta\rceil - 1$

$$|H^{(l)}(s)| \leq_c n^l\ e^{l(S_n-L_n)} \tag{4.39}$$

$$|H^{(l)}(s)| \leq_c n^l e^{(\lceil\beta\rceil-2)(S_n-L_n)} + n\Upsilon(n^2 e^{M_n-L_n}(1-s)^{-1})(1-s)^{-(\lceil\beta\rceil-\beta)} e^{-S_n} e^{(\beta-1)(S_n-L_n)} \tag{4.40}$$

$$\begin{aligned} |H^{(l)}(s)| \leq_c\ & n^l e^{(\lceil\beta\rceil-1)(S_n-L_n)} + n\Upsilon(n^2 e^{M_n-L_n}(1-s)^{-1})\ e^{-S_n} e^{\beta(S_n-L_n)}(1-s)^{-(\lceil\beta\rceil-\beta)} \\ & + n\Upsilon(n^2 e^{M_n-L_n}(1-s)^{-1})\ e^{-S_n} e^{(\beta-1)(S_n-L_n)}(1-s)^{-1-(\lceil\beta\rceil-\beta)} \ a.s. \end{aligned} \tag{4.41}$$

Proof. The Lemma is proved by induction with respect to l. The following relations hold a.s. for every $s \in [0,1)$. For $l = 1$, (4.38) is trivially fulfilled since $f'_{0,n}(1) = e^{S_n}$. First, consider $l < \lceil\beta\rceil - 2$ and assume that (4.38) holds for every $i \le l$. Then it will be proved that (4.39) also holds for l. The induction is completed by proving that (4.38) also holds for $l+1$ if $l+1 < \lceil\beta\rceil - 1$.

By induction assumptions and monotonicity of generating functions and its derivatives, for all $i \le l$ and $s \in [0,1]$,

$$\begin{aligned} f^{(i)}_{k+1,n}(s) \ \le\ f^{(i)}_{k+1,n}(1) \quad &\le_c \quad n^{i-1} e^{S_n}\ e^{(i-1)(S_n - S_k - \min_{j=k,..,n}\{S_j - S_k\})} \\ &\le_c \quad n^{i-1}\ e^{S_n}\ e^{(i-1)(S_n - L_n)}. \end{aligned} \tag{4.42}$$

Lemma A.3.1 in Section A.3 ensures that (see Lemma A.3 for the definition of $u_{j,l}$)

$$\Big|\frac{d^l}{ds^l} h_{k+1}(f_{k+1,n}(s))\Big| \quad = \quad \Big|\sum_{j=1}^{l} h^{(j)}_{k+1}(f_{k+1,n}(s)) u_{j,l}(s)\Big|$$

and by using (4.42)

$$u_{j,l}(s) \quad \le_c \quad n^{l-j}\ e^{jS_n}\ e^{(l-j)(S_n-L_n)} \ \le_c\ n^{l-1}\ e^{S_n}\ e^{(l-1)(S_n-L_n)}\ .$$

By Lemma 4.3.10, for $j < \lceil\beta\rceil - 2$, the derivatives $h^{(j)}_k$ are bounded by a constant that does not depend on Π. Thus

$$\Big|\frac{d^l}{ds^l} h_{k+1}(f_{k+1,n}(s))\Big| \quad \le_c \quad n^{l-1}\ e^{S_n} e^{(l-1)(S_n-L_n)}\ .$$

Then recalling (4.36), we have

$$|H^{(l)}(s)| \quad \le_c \quad \sum_{k=0}^{n-1} n^{l-1}\ e^{-S_k} e^{(l-1)(S_n-L_n)} e^{S_n} \ \le_c\ n^l e^{l(S_n-L_n)}\ ,$$

which results in (4.39) for $l < \lceil\beta\rceil - 2$.

We are now able to prove that (4.38) is fulfilled for $l+1 < \lceil\beta\rceil - 1$. Using Lemma A.3.1 again (see (A.3)) with $f = g_{0,n}$ and $h(x) = 1/x$, we get that

$$\begin{aligned} \frac{d^l}{ds^l} g_{0,n}(s)^{-1} \quad &= \quad \sum_{j=1}^{l} (-1)(-2)\cdots(-j) g_{0,n}(s)^{-(j+1)} u_{j,l}(s) \\ &= \quad -g_{0,n}(s)^{-2} g^{(l)}_{0,n}(s) + \sum_{j=2}^{l} (-1)(-2)\cdots(-j) g_{0,n}(s)^{-(j+1)} u_{j,l}(s), \end{aligned} \tag{4.43}$$

where

$$u_{j,l}(s) \quad = \quad \sum_{i=(i_1,\dots,i_{2j})\in\mathcal{C}(j,l)} c_i (g^{(i_1)}_{0,n}(s))^{i_2} \cdots (g^{(i_{2j-1})}_{0,n})(s))^{i_{2j}}$$

and $\mathcal{C}(j,l) = \big\{(i_1,\dots,i_{2j})) \in \mathbb{N}^{2j} \big| i_1 i_2 + i_3 i_4 + \ldots = l \text{ and } \ i_2 + i_4 + \ldots = j\big\}$. Moreover, the following relation proved in Section 4.3.7 (see (4.48))

$$f^{(l)}(1) \ = \ l g^{(l-1)}(1)$$

together with the induction assumption (i.e. (4.38)) yields for every $i \le l-1$,

$$g^{(i)}_{0,n}(1) \ \le_c\ n^i\ e^{S_n} e^{i(S_n-L_n)}\ .$$

Thus

$$u_{j,l}(1) \ \le_c\ n^l\ e^{jS_n} e^{l(S_n-L_n)}\ .$$

By (4.36), the left-hand side of (4.43) is equal to $(1-s)H^{(l)}(s) - lH^{(l-1)}(s)$. By (4.39), for $l < \lceil\beta\rceil - 2$, $(1-s)H^{(l)}(s)$ vanishes for $s = 1$. Thus letting $s \to 1$ and noting that $g_{0,n}(1) = e^{S_n}$ yields

$$\begin{aligned} g_{0,n}^{(l)}(1) &\leq_c e^{2S_n}\Big(\sum_{j=2}^{l}(-1)(-2)\cdots(-j)e^{-(j+1)S_n}\ n^l\ e^{jS_n}e^{l(S_n-L_n)}\ + l|H^{(l-1)}(1)|\Big) \\ &\leq_c e^{S_n}\ n^l\ e^{l(S_n-L_n)} + e^{2S_n}|H^{(l-1)}(1)|\ . \end{aligned}$$

As we have already proved (4.39) for $l < \lceil\beta\rceil - 2$, we get

$$\begin{aligned} g_{0,n}^{(l)}(1) &\leq_c n^l\ e^{S_n}e^{l(S_n-L_n)} + n^{l-1}\ e^{2S_n}e^{(l-1)(S_n-L_n)} \\ &\leq_c n^l e^{S_n}e^{l(S_n-L_n)}\ . \end{aligned}$$

Using (4.48), we get (4.38) for $l+1$, which completes the induction and proves (4.38) for $l < \lceil\beta\rceil - 1$.

Let us prove the bound on $H^{(l)}(s)$ for $l = \lceil\beta\rceil - 2$. Using again Lemmata 4.3.10 and A.3.1 and (4.38) yields

$$\begin{aligned} &\Big|\frac{d^l}{ds^l}h_{k+1}(f_{k+1,n}(s))\Big| \\ &\quad = \Big|\sum_{j=1}^{l-1}h_{k+1}^{(j)}(f_{k+1,n}(s))u_{j,l}(s) + h_{k+1}^{(l)}(f_{k+1,n}(s))(f'_{k+1,n}(s))^l\Big| \\ &\quad \leq_c n^{l-1}e^{S_n}e^{(\lceil\beta\rceil-3)(S_n-L_n)} + \Upsilon(1/(1-f_{k+1,n}(s)))(1-f_{k+1,n}(s))^{-(\lceil\beta\rceil-\beta)}(f'_{k+1,n}(s))^l\ . \end{aligned} \tag{4.44}$$

By the same arguments as in the proof of Theorem 4.3.4, $\Upsilon(1/(1-f_{k+1,n}(s))) \leq \Upsilon(n^2e^{M_n-L_n}(1-s)^{-1})$ and by convexity, we get

$$(1-f_{k+1,n}(s))^{-(\lceil\beta\rceil-\beta)}\ \leq\ (1-s)^{-(\lceil\beta\rceil-\beta)}(f'_{k+1,n}(s))^{-(\lceil\beta\rceil-\beta)}\ .$$

Using also $f'_{k+1,n}(s) \leq e^{S_n-L_n}$, by (4.44) follows

$$\Big|\frac{d^l}{ds^l}h_{k+1}(f_{k+1,n}(s))\Big| \quad\leq_c\quad n^{l-1}e^{S_n}e^{(\lceil\beta\rceil-3)(S_n-L_n)} + \Upsilon(n^2e^{M_n-L_n}(1-s)^{-1})(1-s)^{-(\lceil\beta\rceil-\beta)}e^{(\beta-2)(S_n-L_n)}\ .$$

Combining this inequality with (4.36) proves (4.40).

This implies that $(1-s)H^{(l)}(s) \to 0$ as $s \to 1$ for $l = \lceil\beta\rceil - 2$. Thus we can apply the same arguments to get an upper bound for $g_{0,n}^{(l)}(1)$ and prove (4.38) for $l = \lceil\beta\rceil - 1$.

Finally, let $l = \lceil\beta\rceil - 1$. We apply just the same arguments as before. Then Lemmata 4.3.10 and A.3.1 yield

$$\begin{aligned} &\Big|\frac{d^l}{ds^l}h_{k+1}(f_{k+1,n}(s))\Big| \\ &\quad = \Big|\sum_{j=1}^{l-2}h_{k+1}^{(j)}(f_{k+1,n}(s))u_{j,l}(s) + lh_{k+1}^{(l-1)}(f_{k+1,n}(s))f_{k+1,n}^{(2)}(s)(f'_{k+1,n}(s))^{l-2} \\ &\qquad\qquad + h_{k+1}^{(l)}(f_{k+1,n}(s))(f'_{k+1,n}(s))^l\Big| \\ &\quad \leq_c n^{l-1}e^{S_n}e^{(l-1)(S_n-L_n)} + \\ &\qquad \Upsilon(n^2e^{M_n-L_n}(1-s)^{-1})\big(e^{(\beta-1)(S_n-L_n)}(1-s)^{-(\lceil\beta\rceil-\beta)} + e^{(\beta-2)(S_n-L_n)}(1-s)^{-1-(\lceil\beta\rceil-\beta)}\big). \end{aligned}$$

Using again (4.36) proves (4.41). □

Proof of Theorem 4.3.6 for $\beta > 2$***.*** Let $l = \lceil \beta \rceil - 1$. W.l.o.g., we assume $\Upsilon \geq 1$. Again, the following relations hold a.s. with respect to the underlying probability measure $\mathbb{P}$. Combining (4.37) and (4.43), we get

$$g_{0,n}^{(l)}(s) \quad = \quad g_{0,n}(1)^2 \Big(-(1-s)H^{(l)}(s) + lH^{(l-1)}(s) + \sum_{j=2}^{l} (-1)(-2)\cdots(-j) g_{0,n}(s)^{-(j+1)} u_{j,l}(s) \Big) \ .$$

Using (4.40), (4.41),and $g_{0,n}(s) \leq e^{S_n}$ for the first terms yields

$$\begin{aligned} g_{0,n}^{(l)}(s) \leq_c\ & e^{2S_n} n^l \Upsilon(n^2 e^{M_n - L_n}(1-s)^{-1}) \Big((1-s) e^{(\lceil \beta \rceil - 1)(S_n - L_n)} + e^{-S_n} e^{\beta(S_n - L_n)} (1-s)^{1-(\lceil \beta \rceil - \beta)} \\ & + e^{-S_n} e^{(\beta-1)(S_n-L_n)} (1-s)^{-(\lceil \beta \rceil - \beta)} + e^{(\lceil \beta \rceil - 2)(S_n - L_n)} + e^{-S_n} e^{(\beta - 1)(S_n - L_n)} (1-s)^{-(\lceil \beta \rceil - \beta)} \Big) \\ & + g_{0,n}^{-(j-1)}(s) u_{j,l}(s). \end{aligned}$$

Using that for every $i \in \mathbb{N}$, $g^{(i)}(s)/(g(s))^i \leq g^{(i)}(1)/(g(1))^i$ (see (4.50), appendix), the definition of $u_{j,l}$, (4.38) and (4.48), we get that

$$g_{0,n}^{-(j-1)}(s) u_{j,l}(s) \leq_c n^{l-1} e^{S_n} e^{(l-1)(S_n - L_n)}.$$

Thus, as $e^{S_n} \leq e^{S_n - L_n}$, we get that

$$\begin{aligned} g_{0,n}^{(l)}(s) \leq_c\ & e^{S_n} n^l \Upsilon(n^2 e^{M_n - L_n}(1-s)^{-1}) \Big((1-s)^{-(\lceil \beta \rceil - \beta)} e^{(\beta - 1)(S_n - L_n)} + (1-s)^{1-(\lceil \beta \rceil - \beta)} e^{\beta(S_n - L_n)} \\ & + (1-s) e^{\lceil \beta \rceil (S_n - L_n)} + e^{(\lceil \beta \rceil - 1)(S_n - L_n)} \Big) + e^{S_n} n^l e^{(\lceil \beta \rceil - 1)(S_n - L_n)}. \end{aligned}$$

As in (4.28), we get the following estimate for every $1/2 \leq s < 1$:

$$g_{0,n}^{(l)}(s) \quad \geq_c \quad s^k \frac{k^{l+1}}{2} \mathbb{P}(Z_n > k | \Pi) \ .$$

Choosing $s = 1 - 1/k$ yields

$$\begin{aligned} \mathbb{P}(Z_n > k|\Pi) \quad \leq_c \quad & e^{S_n} n^l \Upsilon(n^2 e^{M_n - L_n} k) \Big(k^{-\beta} e^{(\beta - 1)(S_n - L_n)} + k^{-(\beta+1)} e^{\beta(S_n - L_n)} \\ & + k^{-\lceil \beta \rceil - 1} e^{\lceil \beta \rceil (S_n - L_n)} + k^{-\lceil \beta \rceil} e^{(\lceil \beta \rceil - 1)(S_n - L_n)} \Big) \ . \end{aligned}$$

Using that the function $x \to a^{-x} \exp((x-1)b)$ is monotone for all $a \geq 1$ and $b \geq 0$, and that $\beta \leq \lceil \beta \rceil < \beta + 1 \leq \lceil \beta \rceil + 1$, we have for all $k \geq 1$,

$$k^{-\lceil \beta \rceil} e^{(\lceil \beta \rceil - 1)(S_n - L_n)} \ \leq \ \max \left\{ k^{-\beta} e^{(\beta - 1)(S_n - L_n)}; k^{-\lceil \beta \rceil - 1} e^{\lceil \beta \rceil (S_n - L_n)} \right\} \ .$$

Combining the two last inequalities leads to

$$P(Z_n > k|\Pi) \quad \leq_c \quad e^{S_n} n^l \Upsilon(n^2 e^{M_n - L_n} k) \max \left\{ k^{-\beta} e^{(\beta-1)(S_n - L_n)}; k^{-\lceil \beta \rceil - 1} e^{\lceil \beta \rceil (S_n - L_n)} \right\} \ ,$$

which completes the proof. □

Proof of the upper bound of Theorem 4.3.1 for $\beta > 2$***.*** The proof now follows the same lines as the proof for $\beta \in (1, 2]$. Theorem 4.3.6 yields

$$\limsup_{n \to \infty} \frac{1}{n} \log \mathbb{P}(Z_n > e^{\theta n}) \quad \leq \quad - \min \left\{ \psi_{\gamma, \beta, \Lambda}(\theta), \psi_{\gamma, \lceil \beta \rceil + 1, \Lambda}(\theta) \right\} \ ,$$

where ψ is defined in (4.16). Using the characterization of ψ proved in the forthcoming Lemma 4.3.8, we deduce that for any $\theta \geq 0$,

$$\psi_{\gamma, \beta, \Lambda}(\theta) \ \leq \ \psi_{\gamma, \lceil \beta \rceil + 1, \Lambda}(\theta) \ .$$

Thus

$$\min \left\{ \psi_{\gamma, \beta, \Lambda}(\theta), \psi_{\gamma, \lceil \beta \rceil + 1, \Lambda}(\theta) \right\} \ = \ \psi_{\gamma, \beta, \Lambda}(\theta) \ = \ \psi(\theta)$$

and we get the expected upper bound. □

4.3.5 Proof of Theorem 4.3.2

By assumption, there exists a constant $d < \infty$ such that for every $\beta > 0$,

$$\mathcal{P}(R > z | R > 0) \quad \leq \quad d(m \wedge 1) z^{-\beta} \quad \text{a.s.}$$

Then we can apply the upper bound in Theorem 4.3.1 for every $\beta > 0$. This yields for all $\beta > 0$ and $\theta \geq 0$,

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}(Z_n > e^{\theta n}) \quad \leq \quad -\psi_{\gamma,\beta,\Lambda}(\theta) \ .$$

Taking the limit $\beta \to \infty$, the monotone convergence of $\psi_{\gamma,\beta,\Lambda}$ yields

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}(Z_n > e^{\theta n}) \quad \leq \quad -\psi_{\gamma,\infty,\Lambda}(\theta) \ ,$$

where

$$\begin{aligned}
\psi_{\gamma,\infty,\Lambda}(\theta) \quad &:= \quad \lim_{\beta\to\infty} \inf_{t\in[0,1], s\in[0,\theta]} \Big\{ t\gamma + \beta s + (1-t)\Lambda((\theta - s)/(1-t)) \Big\} \\
&= \quad \inf_{t\in[0,1]} \Big\{ t\gamma + (1-t)\Lambda(\theta/(1-t)) \Big\} (= \Gamma(\theta)) \ .
\end{aligned}$$

This yields the upper bound and the lower bound follows readily the proof presented in Section 4.3.2 where the natural associated path is considered (or see Section 4.2). □

4.3.6 Characterization of the rate function ψ

Here we state a characterization of ψ similar to Lemma 4.2.1.

Lemma 4.3.8. *Let $0 \leq \gamma \leq \Lambda(0)$ and $\beta > 0$. The rate function ψ defined for $\theta \geq 0$ by*

$$\psi(\theta) = \inf_{t\in[0,1], s\in[0,\theta]} \{ t\gamma + \beta s + (1-t)\Lambda((\theta - s)/(1-t)) \} \tag{4.45}$$

is the largest convex function such that for all $x, \theta \geq 0$

$$\psi(0) = \gamma, \qquad \psi(\theta + x) \leq \psi(\theta) + \beta x, \qquad \psi(\theta) \leq \Lambda(\theta) \ . \tag{4.46}$$

Proof. We denote the infimum of (4.45) by ι. First, we prove that ι is convex. Using convexity of Λ, for any $\theta', \theta'' \geq 0$ and $\epsilon > 0$ there exist $t', t'' \in (0,1]$ and $s' \in [0, \theta']$, $s'' \in [0, \theta'']$, such that for every $\lambda \in [0,1]$

$$\begin{aligned}
&\lambda\iota(\theta') + (1-\lambda)\iota(\theta'') \\
&\geq \ \lambda\big(t'\gamma + \beta s' + (1-t')\Lambda((\theta' - s')/(1-t'))\big) \\
&\qquad + (1-\lambda)\big(t''\gamma + \beta s'' + (1-t'')\Lambda((\theta'' - s'')/1 - t''))\big) - \epsilon \\
&= \ \Big(\lambda(1-t') + (1-\lambda)(1-t'')\Big)\gamma + \Big(\lambda s' + (1-\lambda)s''\Big)\beta \\
&\qquad + (\lambda t' + (1-\lambda)(1-t''))\frac{\lambda(1-t')}{\lambda(1-t') + (1-\lambda)(1-t'')}\Lambda((\theta' - s')/(1-t')) \\
&\qquad + (\lambda(1-t') + (1-\lambda)(1-t''))\frac{(1-\lambda)(1-t'')}{\lambda(1-t') + (1-\lambda)(1-t'')}\Lambda((\theta'' - s'')/(1-t'')) - \epsilon \\
&\geq \ \Big(1 - (\lambda(1-t') + (1-\lambda)(1-t''))\Big)\gamma + \Big(\lambda s' + (1-\lambda)s''\Big)\beta \\
&\qquad + (\lambda(1-t') + (1-\lambda)(1-t''))\Lambda\Big(\frac{\lambda(\theta' - s') + (1-\lambda)(\theta'' - s'')}{\lambda(1-t') + (1-\lambda)(1-t'')}\Big) - \epsilon \\
&\geq \ \iota(\lambda\theta' + (1-\lambda)\theta'') - \epsilon \ .
\end{aligned}$$

Letting $\epsilon \to 0$ entails that ψ is convex.

Second, following the previous computation, we verify that ι fulfills (4.46). For any $\theta \geq 0$ and $\epsilon > 0$, there exist $t' \in [0,1)$ and $s' \in [0,\theta]$ such that

$$\begin{aligned} \iota(\theta) &\geq t'\gamma + \beta s' + (1-t')\Lambda\big((\theta - s')/(1-t')\big) - \epsilon \\ &= t'\gamma + \beta(s'+x) + (1-t')\Lambda\big((\theta + x - (s'+x))/(1-t')\big) - \beta x - \epsilon \\ &\geq \inf_{t\in[0,1],\tilde{s}\in[0,\theta+x]} \big\{t\gamma + \beta\tilde{s} + (1-t)\Lambda((\theta + x - \tilde{s})/(1-t))\big\} - \beta x - \epsilon \ . \end{aligned}$$

Taking the limit $\epsilon \to 0$ yields the second property in (4.46). Furthermore, setting $t = 0$ and $s = 0$ implies $\psi(\theta) \leq \Lambda(\theta)$ and $t \to 1$ entails that $\psi(0) \leq \gamma$. This completes the proof of (4.46).

Finally, let κ be any convex function which satisfies (4.46). Using these assumptions ensures that for all $t \in [0,1)$ and $0 \leq s \leq \theta$,

$$\begin{aligned} t\gamma + \beta s + (1-t)\Lambda((\theta - s)/(1-t)) &\geq t\kappa(0) + \beta s + (1-t)\kappa((\theta - s)/(1-t)) \\ &\geq \beta s + \kappa\big(t0 + (1-t)(\theta - s)/(1-t))\big) \\ &= \beta s + \kappa(\theta - s) \\ &\geq \kappa(\theta) \ . \end{aligned}$$

By taking the infimum over s and t we get $\iota(\theta) \geq \kappa(\theta)$ and the proof is complete. □

Finally, let κ be any convex function which satisfies (4.46). Using the properties in (4.46) ensures that for all $t \in [0,1)$ and $0 \leq s \leq \theta$,

$$\begin{aligned} t\gamma + \beta s + (1-t)\Lambda((\theta - s)/(1-t)) &\geq t\kappa(0) + \beta s + (1-t)\kappa((\theta - s)/(1-t)) \\ &\geq \beta s + \kappa\big(t0 + (1-t)(\theta - s)/(1-t)\big) \\ &= \beta s + \kappa(\theta - s) \\ &\geq \kappa(\theta) \ . \end{aligned}$$

Taking the infimum over s and t, we get $\psi(\theta) \geq \kappa(\theta)$ and the proof is complete. □

Finally, another characterization of ψ results from Lemma 4.3.8 (see Figure 4.5): Let θ^* and $\theta^\dagger$ be defined as in (4.8) and (4.18) and assume $0 < \theta^* < \theta^\dagger < \infty$. As a convex and monotone function, Λ has at most one jump (to infinity). Let the first jump be in $0 < \theta_j < \infty$. If there is no jump, we set $\theta_j = \infty$. For $\theta < \theta_j$, $\Lambda(\theta)$ is differentiable. As Λ is also continuous from below, $\Lambda(\theta_j) < \infty$. We will use the continuity to define another characterization of ψ.

By the preceding characterization, ψ is the largest convex function, starting in $\psi(0) = \gamma$, being at most as large as Λ and having at most slope β.

The largest convex function through the point $(0,\gamma)$ being smaller than or equal to Λ has to be linear and has to be a tangent of Λ. By definition of θ^*, the tangent of Λ in θ^* touches the point $(0,\gamma)$. Thus ψ is linear for $\theta < \theta^*$ and follows this tangent. For $\theta > \theta^*$, ψ is identical to Λ until the slope of Λ is exactly β (or until Λ jumps to infinity). At this point $\theta^\dagger$, the last condition becomes important and ψ is linear with slope β for $\theta > \theta^\dagger$. Summing up,

$$\psi(\theta) = \begin{cases} \gamma(1 - \frac{\theta}{\theta^*}) + \frac{\theta}{\theta^*}\Lambda(\theta^*) & , \text{ if } \theta \leq \theta^* \\ \Lambda(\theta) & , \text{ if } \theta^* < \theta < \theta^\dagger \\ \beta(\theta - \theta^\dagger) + \Lambda(\theta^\dagger) & , \text{ if } \theta \geq \theta^\dagger \end{cases} \ .$$

If $\Lambda'(0) > \beta$, then $\theta^\dagger = 0$ and $\psi(\theta) = \gamma + \beta\theta$. If $\gamma = \Lambda(0)$, then $\theta^* = 0$. We refrain from describing other degenerated cases.

4.3.7 Bounds for generating functions

Let R be a random variable with values in $\{0, 1, 2, ...\}$, expectation m, distribution $(p_k)_{k\in\mathbb{N}}$, and generating function f. Let us define

$$q_k \ := \ \mathcal{P}(R > k)$$

and the following function associated with f,

$$g(s) \ := \ \sum_{k=0}^{\infty} s^k q_k \ = \ \frac{1-f(s)}{1-s} \ , \tag{4.47}$$

where the last identity comes from Cauchy product of power series (see also Section 4.2.1). We recall that the l^{th} derivative of a function f is denoted by $f^{(l)}$ and that $f^{(l)}(s)$ and $g^{(l)}(s)$ exist for every $s \in [0, 1)$. As

$$f^{(l)}(s) \ = \ \sum_{k=0}^{\infty} k(k-1)\cdots(k-l+1)s^{k-l}p_k, \qquad g^{(l)}(s) = \sum_{k=0}^{\infty} k(k-1)\cdots(k-l+1)s^{k-l}q_k,$$

all derivatives of f and g are nonnegative, nondecreasing functions. We are using g instead of f in the proofs since the associated sequence $(q_k)_{k\in\mathbb{N}}$ is monotone, which is more convenient. A straightforward calculation yields

$$f^{(l)}(1) \quad = \quad lg^{(l-1)}(1). \tag{4.48}$$

Thus $g^{(l-1)}(1)$ and $f^{(l)}(1)$ both essentially describe the l-th moment of the corresponding probability distribution. Calculating the l-th derivative of $f(s) = 1-(1-s)g(s)$ yields

$$f^{(l)}(s) \quad = \quad lg^{(l-1)}(s) - (1-s)g^{(l)}(s). \tag{4.49}$$

For the proofs, we will use g instead of f since the associated sequence $(q_k)_{k\in\mathbb{N}}$ is monotone, which is more convenient. Next, we prove that for every $i \in \mathbb{N}$,

$$g^{(i)}(s) \cdot (g(1))^i \leq (g(s))^i \cdot g^{(i)}(1). \tag{4.50}$$

We will prove the result by induction. For $i = 1$, define a random variable Y with distribution $(q_k/g(1))_{k\in\mathbb{N}_0}$. Then

$$g'(s) \cdot g(1) = \mathbb{E}[s^Y Y]g(1)^2 \leq \mathbb{E}[s^Y]\mathbb{E}[Y]g(1)^2 = g(s)g'(1). \tag{4.51}$$

as s^Y and Y are obviously negatively correlated for $s \in [0, 1]$. Note that the above inequality remains true if g' is replaced by $g^{(i)}$, $i \in \mathbb{N}$. Next, let (4.50) be fulfilled for i. Thus, using (4.49) and the induction assumption and monotonicity of $g^{(i)}$ yield

$$\begin{aligned}
\frac{g^{(i+1)}(s)}{(g(s))^{i+1}} &= \frac{(i+1)g^{(i)}(s) - f^{(i+1)}(s)}{(1-s)(g(s))^{i+1}} \leq \frac{(i+1)g^{(i)}(1)}{(1-s)(g(1))^i g(s)} - \frac{f^{(i+1)}(s)}{(1-s)(g(1))^i g(s)} \\
&= \frac{f^{(i+1)}(1) - f^{(i+1)}(s)}{(1-s)(g(1))^i g(s)} = \frac{1}{(g(1))^i g(s)} \sum_{k=0}^{\infty} k(k-1)\cdots(k-i)\frac{1-s^{k-i-1}}{1-s} p_k \\
&= \frac{g^{(i+1)}(s)}{g(1))^i g(s)} \leq \frac{g^{(i+1)}(1)}{g(1))^{i+1}}.
\end{aligned}$$

In the last step, we used (4.51) with g' replaced by $g^{(i+1)}$.

For $\mu \in (0, 1]$, we also define the function

$$h_\mu(s) \ := \ \frac{g(1)^\mu - g(s)^\mu}{(g(1)g(s)(1-s))^\mu} \ . \tag{4.52}$$

The following useful lemmata describe versions of Assumption $\mathcal{H}(\beta)$ in terms of the function h_μ. Noting that $g(0) = q_0 = \mathbb{P}(L > 0|f)$ and $g(1) = m$, we can rewrite Assumption $\mathcal{H}(\beta)$ in the following way

$$q_k \ \leq \ d \cdot g(0) \cdot (g(1) \wedge 1) \cdot k^{-\beta} \qquad (k \geq 1) \ . \tag{4.53}$$

Lemma 4.3.9. *Let $\beta > 1$ and assume that (4.53) holds for some constant $0 < d < \infty$. Then for every $0 < \mu < (\beta - 1) \wedge 1$, there exists a constant $c = c(\beta, d, \mu)$ such that for every $s \in [0,1]$,*

$$h_\mu(s) \quad \le \quad c \; . \tag{4.54}$$

The above bound also holds for $\mu = 1$ if $\beta > 2$. Moreover, if $\beta \in (1,2]$, there exists a nondecreasing positive slowly varying function $\Upsilon = \Upsilon(\beta, d)$ such that for every $s \in [0,1)$,

$$h_{\beta-1}(s) \quad \le \quad \Upsilon(1/(1-s)) \tag{4.55}$$
$$-h'_{\beta-1}(s) \quad \le \quad \Upsilon(1/(1-s))/(1-s) \; . \tag{4.56}$$

Note that Υ depends on L (or g) only through the values of d and β. Then under Assumption $\mathcal{H}(\beta)$, we derive from this lemma a nonrandom constant bound.

In the proofs, we use the notation $\le_c$ again, which means that the inequality is fulfilled up to a multiplicative constant which depends on β and μ but is independent of s and the order of the differentiation.

Proof. Using $g(s) \ge g(0)$ yields

$$\begin{aligned} h_\mu(s) &= \frac{g(1)^\mu - g(s)^\mu}{(g(1)g(s)(1-s))^\mu} \\ &\le \frac{g(1)^\mu - g(s)^\mu}{(g(1)g(0)(1-s))^\mu} \\ &\le (g(1) \wedge 1)^{-1} \frac{(\sum_{k=0}^\infty g(0)^{-1} q_k)^\mu - (\sum_{k=0}^\infty s^k q_k g(0)^{-1})^\mu}{(1-s)^\mu} \; . \end{aligned} \tag{4.57}$$

Since $\mu \in (0,1]$, the function $x \to x^\mu$ is concave, such that $a^\mu - x^\mu \le \mu x^{\mu-1}(a-x)$ for all $0 \le x \le a$. Moreover

$$1 \; = \; q_0/g(0) \; \le \; x \; := \; \sum_{k=0}^\infty s^k q_k g(0)^{-1} \; \le \; a \; := \; \sum_{k=0}^\infty q_k g(0)^{-1} \; . \tag{4.58}$$

Then $x^{\mu-1} \le 1$ and using the inequality of concavity in (4.57) with $q_k \le d g(0) \cdot (g(1) \wedge 1) \cdot k^{-\beta}$ leads to

$$\begin{aligned} h_\mu(s) &\le \mu (g(1) \wedge 1)^{-1} x^{\mu-1} \frac{\sum_{k=0}^\infty g(0)^{-1} q_k [1 - s^k]}{(1-s)^\mu} \\ &\le_c \frac{\sum_{k=1}^\infty (1-s^k) k^{-\beta}}{(1-s)^\mu} \\ &= (1-s)^{1-\mu} \sum_{k=1}^\infty \frac{1-s^k}{1-s} k^{-\beta} \\ &= (1-s)^{1-\mu} \sum_{k=1}^\infty k^{-\beta} \sum_{j=0}^{k-1} s^j \\ &= (1-s)^{1-\mu} \sum_{j=0}^\infty s^j \sum_{k=j+1}^\infty k^{-\beta} \\ &\le_c (1-s)^{1-\mu} \sum_{j=0}^\infty s^j (j+1)^{-\beta+1} \; . \end{aligned}$$

The estimates (4.54) and (4.55) on h_μ for $0 < \mu < (\beta-1) \wedge 1$ and $\mu = \beta - 1$ now follow directly from (A.2.2). For $\mu = 1$, $\beta > 2$ and $s = 1$, the sum is finite and (4.54) also holds in this case.

For the second part of the lemma, we explicitly calculate the first derivative of $h_{\beta-1}$ by using the formula

$$h_{\beta-1}(s) g(s)^{\beta-1} \quad = \quad \frac{g(1)^{\beta-1} - g(s)^{\beta-1}}{g(1)^{\beta-1}(1-s)^{\beta-1}} \; .$$

Differentiating both sides yields

$$h'_{\beta-1}(s)g(s)^{\beta-1} + (\beta-1)h_{\beta-1}(s)g(s)^{\beta-2}g'(s) \;=\; \frac{(\beta-1)([g(1)^{\beta-1} - g(s)^{\beta-1}] - (1-s)g(s)^{\beta-2}g'(s))}{g(1)^{\beta-1}(1-s)^{\beta}}$$

and thus

$$-h'_{\beta-1}(s) \;=\; (\beta-1)\Big(\frac{h_{\beta-1}(s)g'(s)}{g(s)} + \frac{g'(s)}{g(s)g(1)^{\beta-1}(1-s)^{\beta-1}} - \frac{g(1)^{\beta-1} - g(s)^{\beta-1}}{g(s)^{\beta-1}g(1)^{\beta-1}(1-s)^{\beta}}\Big)\;.$$

As g is nondecreasing, we can skip the last term which is negative. Using (4.53) and (4.55), we get

$$-h'_{\beta-1}(s) \;\le_c\; \frac{g(0)\cdot(g(1)\wedge 1)\cdot\sum_{k=1}^{\infty} ks^{k-1}k^{-\beta}}{g(s)}\Big(\Upsilon(1/(1-s)) + \frac{1}{g(1)^{\beta-1}(1-s)^{\beta-1}}\Big)\;.$$

Moreover, $g(s) \ge g(0)$ and $g(1)^{-(\beta-1)}\cdot(g(1)\wedge 1) \le 1$ for $\beta - 1 \in (0,1]$, so

$$-h'_{\beta-1}(s) \;\le_c\; \sum_{k=1}^{\infty} s^{k-1}k^{-\beta+1}\Big(\Upsilon(1/(1-s)) + \frac{1}{(1-s)^{\beta-1}}\Big)\;.$$

The result now follows from (A.2.2) and the fact that the product of two slowly varying functions is still slowly varying. □

In the second part of this section, the function

$$h(s) \;=\; h_1(s) \;=\; \frac{g(1)-g(s)}{g(1)g(s)(1-s)}$$

is considered.

Lemma 4.3.10. *We assume that (4.53) holds for some $\beta > 1$. Then there exists a constant $c = c(\beta, d) < \infty$ such that for every $s \in [0,1)$,*

$$\begin{aligned} |h^{(l)}(s)| &\le c && \text{if } 0 \le l < \beta - 2 \\ |h^{(\lceil\beta\rceil-2)}(s)| &\le c\Upsilon(1/(1-s))\;(1-s)^{-(\lceil\beta\rceil-\beta)} && \text{if } \beta \ge 2 \\ |h^{(\lceil\beta\rceil-1)}(s)| &\le c\;\Upsilon(1/(1-s))\;(1-s)^{-1-(\lceil\beta\rceil-\beta)}\;. \end{aligned} \tag{4.59}$$

Proof. By (4.52) and Cauchy product of power series, for every $s \in [0,1)$,

$$g(s)g(1)h(s) \;=\; \frac{g(1)-g(s)}{1-s} \;=\; \sum_{k=0}^{\infty} s^k(q_{k+1} + q_{k+2} + \ldots)\;.$$

Thus, the l-th derivative of $g(s)h(s)$ is

$$\sum_{j=0}^{l}\binom{l}{j} g^{(j)}(s)h^{(l-j)}(s) \;=\; g(1)^{-1}\sum_{k=0}^{\infty} k(k-1)\cdots(k-l+1)s^{k-l}(q_{k+1} + q_{k+2} + \ldots)\;.$$

Moreover, (4.53) ensures that for all $s \in [0,1)$ and $j < \beta - 2$,

$$g^{(j)}(s) \;\le\; g^{(j)}(1) \;\le\; \sum_{k=0}^{\infty} k^j q_k \;\le_c\; g(0)(g(1)\wedge 1)\;.$$

Combining the last two expressions and using $g(s)^{-1} \le g(0)^{-1}$ yields

$$\begin{aligned} |h^{(l)}(s)| &\le_c\; g(s)^{-1}\Big(g(1)^{-1}g(0)\cdot(g(1)\wedge 1)\cdot\sum_{k=0}^{\infty} k^l s^{k-l}\sum_{j=k+1}^{\infty} j^{-\beta} + \sum_{j=1}^{l}\binom{l}{j} g^{(j)}(1)|h^{(l-j)}(s)|\Big) \\ &\le_c\; \sum_{k=0}^{\infty} k^l s^{k-l}\sum_{j=k+1}^{\infty} j^{-\beta} + \sum_{j=1}^{l}\binom{l}{j}|h^{(l-j)}(s)|\;. \end{aligned} \tag{4.60}$$

The first statement of the lemma is proved by induction on l. For $l = 0$, it is already included in Lemma 4.3.9. Assuming that the bound holds for $l' < l < \beta - 2$, the previous inequality ensures that

$$|h^{(l)}(s)| \ \leq_c \ 1 \ + \sum_{j=0}^{l-1} |h^{(j)}(s)|$$

since $\sum_{k=0}^{\infty} k^l \sum_{j=k+1}^{\infty} j^{-\beta} < \infty$. This ends up the induction and proves the first estimate in (4.59).

Next $l = \lceil \beta \rceil - 2$ is considered and 'the induction is continued'. Using the bound of $h^{(l)}$ for $l < \beta - 2$ and (4.60) yields

$$|h^{(l)}(s)| \ \leq_c \ \sum_{k=0}^{\infty} k^l s^k \sum_{j=k+1}^{\infty} j^{-\beta} + 1 \ \leq_c \ \sum_{k=1}^{\infty} s^k k^{\lceil \beta \rceil - 2} k^{-\beta+1} + 1 \ \leq_c \ \sum_{k=1}^{\infty} s^k k^{-(1-(\lceil \beta \rceil - \beta))} \ .$$

Then the second estimate of the lemma follows from (A.2.2).

Finally, the bound for $l = \lceil \beta \rceil - 1$ is proved in the same way. By (4.60):

$$\begin{aligned} |h^{(l)}(s)| \ &\leq_c \ \sum_{k=0}^{\infty} k^l s^k \sum_{j=k+1}^{\infty} j^{-\beta} + l g^{(2)}(1) |h^{(l-1)}(s)| + 1 \\ &\leq_c \ \sum_{k=1}^{\infty} s^k k^{\lceil \beta \rceil - \beta} + \sum_{k=1}^{\infty} s^k k^{-(1-(\lceil \beta \rceil - \beta))} + 1 \\ &\leq_c \ \sum_{k=1}^{\infty} s^k k^{\lceil \beta \rceil - \beta} \end{aligned}$$

and (A.2.2) yields the claim. □

4.4 Lower deviations: A result for geometric offspring distributions

4.4.1 Main result

The lower deviations for supercritical BPREs without extinction (i.e. each individual having at least one offspring) have been studied in [BB09]. It turns out that a phase transition occurs. Here, this result is generalized for supercritical BPREs when the offspring distributions are of linear fractional[22] (and extinction is allowed). Then direct calculations with generating functions are feasible. A corresponding result for supercritical BPREs with more general offspring distributions is an open problem.

The first assumption constricts the set of possible offspring distributions. The generating function of them has to be of linear fractional form.

Assumption 4.3. *Q takes $\mathbb{P}$–a.s. values in the set of all probability measures $\mathcal{A} \subset \Delta$ with the following property:*
For $\mathcal{P} \in \mathcal{A}$, let R be a random variable with distribution $\mathcal{P}$ and for $s \in [0, 1]$

$$f_R(s) \ := \ \mathcal{E}[s^R] \ = \ 1 - \frac{1-s}{m_R^{-1} + 1/2 \ b_R m_R^{-2}(1-s)} \ ,$$

where $m_R = \sum_{k=0}^{\infty} k\mathcal{P}(R = k)$ and $b_R = \sum_{k=0}^{\infty} k(k-1)\mathcal{P}(R = k)$. There are constants $0 < c_1 < c_2 < \infty$ (uniformly for all $\mathcal{P} \in \mathcal{A}$) such that

$$c_1 \ < \ 1/2 \ b_R m_R^{-2} \ < \ c_2 \ .$$

[22] i.e. have offsprings with generating functions that are linear fractional

This assumption is fulfilled e.g. for geometric offspring distributions. In this case, $1/2\ b_R m_R^{-2} = 1$. All distributions fulfilling Assumption 4.3 have geometric tails, but the second free parameter b_R allows to change the probability of the event $\mathcal{P}(R = 0)$.

Under Assumption 4.3, we can explicitely calculate the probability weights of Z_n, conditioned on the environment. Define

$$\begin{aligned} U_n &:= e^{-S_n} , \\ V_n &:= \sum_{k=0}^{n-1} 1/2\ b_k m_k^{-2} U_j . \end{aligned}$$

Then for all $z \in \mathbb{N}_0$, (see [Koz06, p. 156])

$$\mathbb{P}(Z_n > z|\Pi) = \frac{1}{(U_n + V_n)(1 + U_n/V_n)^z}$$

and thus

$$\mathbb{P}(1 \leq Z_n \leq z) = (U_n + V_n)^{-1}\big(1 - (1 + U_n/V_n)^{-z}\big) . \tag{4.61}$$

For convenience, we assume that the moment generating function is finite everywhere and that, with a positive probability, there are also unfavorable offspring distributions.

Assumption 4.4. *For all $s \in \mathbb{R}$,*

$$\varphi(s) := \mathbb{E}[e^{sX}] < \infty$$

and $\mathbb{P}(X < 0) > 0$.

Again, we require the rate function $\Lambda : \mathbb{R} \to \bar{\mathbb{R}}^+$ which fulfills

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}(S_n \geq \theta n) = -\Lambda(\theta) \quad \text{for all } \theta \geq \mathbb{E}[X] \tag{4.62}$$

and

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}(S_n \leq \theta n) = -\Lambda(\theta) \quad \text{for all } \theta < \mathbb{E}[X] .$$

It is a convex function, defined by

$$\Lambda(\theta) = \sup_{s\in\mathbb{R}} \big\{ s\theta - \log \varphi(s) \big\}$$

with $\varphi(s) = \mathbb{E}[e^{sX}]$. In contrast to Definition 4.4, the supremum is taken over all $s \in \mathbb{R}$, as we are here also interested in lower deviations. Λ is decreasing in $(-\infty, \mathbb{E}[X])$ and increasing in $(\mathbb{E}[X], \infty)$.

Remark. *The assumption $\mathbb{P}(X < 0) > 0$ assures that $\Lambda(0) < \infty$, i.e. exceptionally small values may also be realized by an exceptional environment. We refrain from proving that in the case $X > 0$ $\mathbb{P}$–a.s. (i.e. $\Lambda(0) = \infty$) , the following Theorem 4.4.1 is still correct, but the interpretation may be different. E.g. in the case $X = const.$ $\mathbb{P}$–a.s., exceptional values may only be realized by the branching mechanism, which is similar to a classical Galton-Watson process. The proof is slightly more involved as for the representation of the rate function in Lemma 4.4.4, θ^* has to be defined differently.*

Define

$$\varrho := -\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}(Z_n = 1) . \tag{4.63}$$

The existence and characterization of this limit will be proved later. Then the rate function χ for Z is defined by

$$\chi(\theta) = \inf_{t\in[0,1)} \big\{ t\varrho + (1-t)\Lambda(\theta/(1-t)) \big\} . \tag{4.64}$$

Our main result is the following theorem:

Theorem 4.4.1. *Under Assumptions 4.3 and 4.4, for all* $0 < \theta < \mathbb{E}[X]$

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}(1 \le Z_n \le e^{\theta n}) \quad = \quad -\chi(\theta)$$

and for all $\theta \ge \mathbb{E}[X]$,

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}(Z_n \ge e^{\theta n}) \quad = \quad -\chi(\theta) \ (= \Lambda(\theta)) \ .$$

The second statement of the Theorem results from Theorem 4.3.2.

Open Problem. *The proof of Theorem 4.4.1 under more general conditions is an open problem. Namely, for some constant* $c > 0$, *let*

$$\varrho \quad := \quad -\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}(1 \le Z_n \le c) \ . \tag{4.65}$$

We claim here that the (4.64) is still the correct rate function for general offspring distributions. The existence of the above limit in $[0, \infty]$ *can be proved using superadditivity arguments. Provided it is finite, the problem is to find an appropriate upper bound for* $\mathbb{P}(1 \le Z_n \le e^{\theta n})$. *Some condition on the tails of the offspring distributions has to be specified. In the linear fractional case, as we will see below,*

$$\varrho \ = \ -\log \mathbb{E}\Big[f'\big(\mathbb{P}(\exists n \in \mathbb{N} \text{ with } Z_n = 0 | \Pi = (f, f, f, \ldots))\big)\Big] \ . \tag{4.66}$$

Define $p_{ex,f} := \mathbb{P}(\exists n \in \mathbb{N} \text{ with } Z_n = 0 | \Pi = (f, f, f, \ldots)$. *As it is well-known for Galton-Watson processes (see [AN72]),* $p_{ex,f}$ *is the smallest fixpoint of the generating function* f, *i.e.* $p_{ex,f} = \inf\{s > 0 | f(s) = s\}$. *Thus in the linear fractional case, it is easy to calculate the above probability and* $\varrho = -\log \mathbb{E}[e^{-X}]$.
It is an open problem to generalize the relation (4.66).

Interpretation of the rate function

Consider the large deviation event $\{1 \le Z_n \le e^{\theta n}\}$ for some $0 < \theta < \mathbb{E}[X]$ and n large, that is we observe a population being much smaller than expected, but still alive. A possible path that led to this event looks as follows (see Figure 4.7).

During a first period, until generation $\lfloor tn \rfloor$ $(0 \le t \le 1)$, the population stays small within a good environment (i.e. the associated random walk grows linearly). This has exponentially small probability of order $e^{-\varrho\lfloor nt\rfloor + o(n)}$. Later, the population grows in a – compared to the typical – less favorable environment, i.e. $\{S_n - S_{\lfloor nt \rfloor} \le \theta n\}$. This atypical environment sequence has also exponentially small probability, of order $e^{-\Lambda(\theta/(1-t))\lfloor n(1-t)\rfloor + o(n)}$. The probability of the large deviation event is then results from maximizing the product of these two probabilities.

By Lemma 4.2.1, χ is also characterized as the largest convex function such that

$$\begin{array}{rcll} \chi(0) & \le & -\log \mathbb{E}\big[e^{-X}\big] & \\ \chi(\theta) & \le & \Lambda(\theta) & , \quad \forall \theta \ge 0 \ . \end{array} \tag{4.67}$$

From the above representation, the shape of χ can already be guessed. For the formal proof, see Section 4.4.2. Define

$$\theta^\star \quad := \quad \frac{\mathbb{E}\big[Xe^{-X}\big]}{\mathbb{E}\big[e^{-X}\big]} \ .$$

Then the function χ is also given by

$$\chi(\theta) \quad = \quad \begin{cases} -\theta - \log \mathbb{E}\big[e^{-X}\big] & , \ \text{if } \theta \le \theta^\star \\ \Lambda(\theta) & , \ \text{if } \theta > \theta^\star \end{cases} \ . \tag{4.68}$$

Note that $\Lambda'(\theta^\star) = -1$. Figure 4.8 illustrates the shape of χ in the case $\Lambda(0) > -\log \mathbb{E}\big[e^{-X}\big]$.

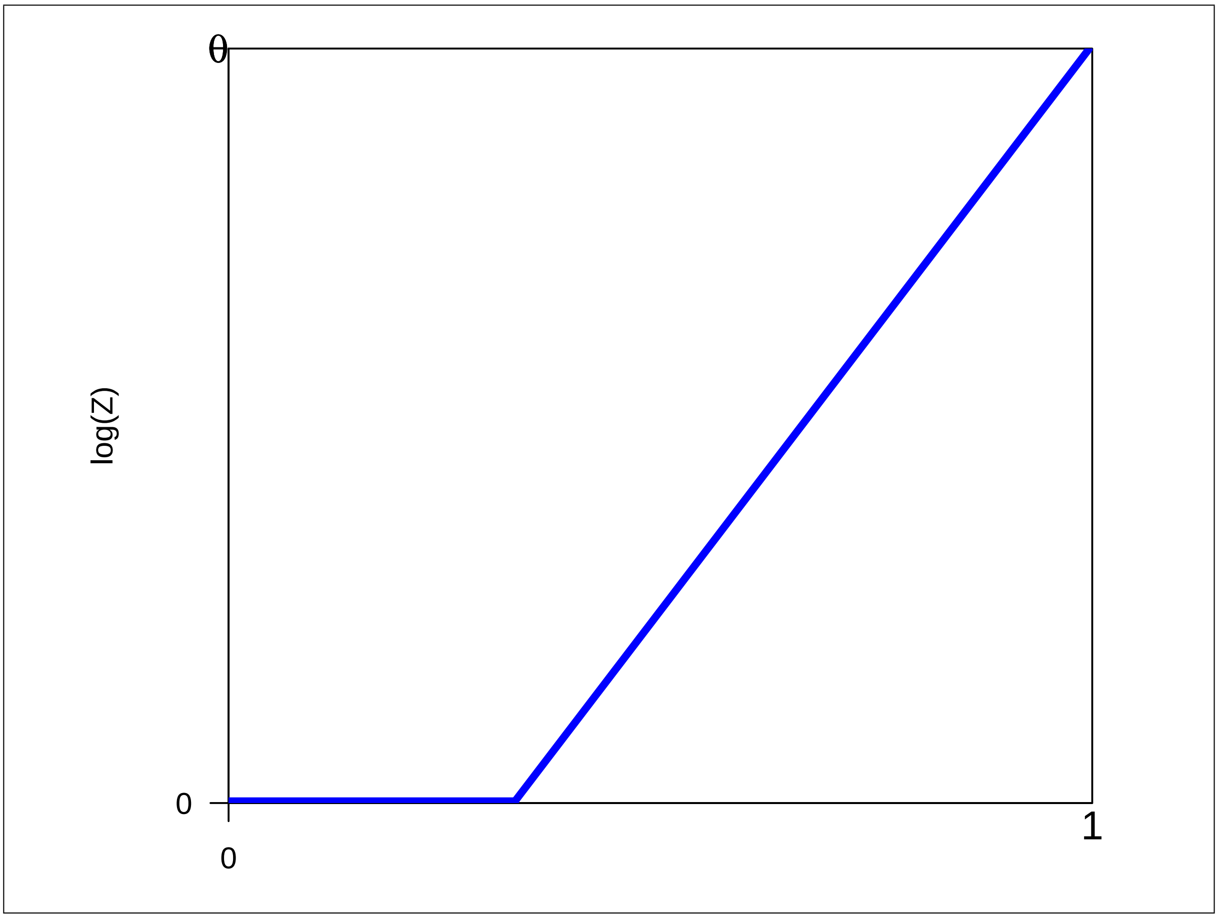


Figure 4.7: Most probable path for the event $\{1 \le Z_n \le e^{\theta n}\}$ with $\theta < \theta^*$.

4.4.2 Proof of Theorem 4.4.1

First note that, conditioned on the associated random walk having exceptionally small values, the same is true for the branching process (compare e.g. [BB09, Proposition 1]):

Lemma 4.4.2. *Under Assumption 4.3 and for all* $\theta > 0$

$$\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}(1 \le Z_n \le e^{\theta n} | S_n \le \theta n) \quad = \quad 0 \ .$$

The result is covered by Lemma 5 in the proof of Proposition 1 in [BB09].

Next, a characterization of the probability that there is just one individual is required:

Lemma 4.4.3. *Under Assumption 4.3 and if* $\mathbb{E}[X] > 0$*, the limit*

$$\varrho \quad := \quad - \lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}(Z_n = 1)$$

exists and

$$\varrho \quad = \quad \begin{cases} -\log \mathbb{E}\big[e^{-X}\big] & , \quad \textit{if} \ \ \mathbb{E}[Xe^{-X}] \ge 0 \\ \Lambda(0) & , \quad \textit{else} \end{cases} \quad . \tag{4.69}$$

Proof. As the offspring distributions are linear fractional, we can explicitely calculate the desired probability (see (4.61)):

$$\mathbb{P}(Z_n = 1|\Pi) \quad = \quad \frac{U_n}{V_n^2(1 + U_n/V_n)^2} \ . \tag{4.70}$$

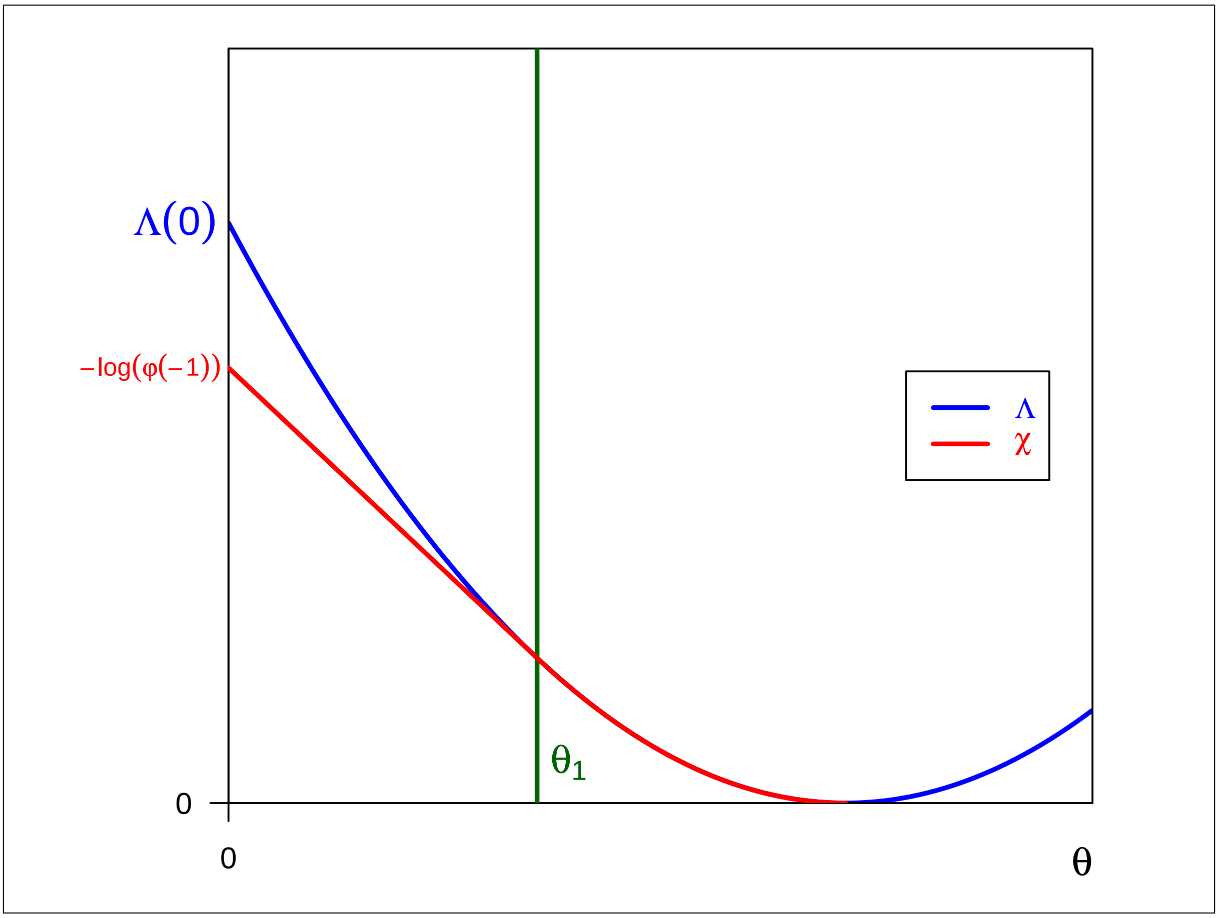


Figure 4.8: χ and Λ with $\theta_1 = \theta^\star$.

By Assumption 4.3 and the definition of V_n,

$$(c_1 \ \leq) \ c_1 e^{-L_n} \quad \leq \quad V_n \ \leq \ c_2 \ n \ e^{-L_n} \ ,$$

and thus

$$c_1^{-1} e^{-(S_n - L_n)} \quad \geq \quad U_n/V_n \ \geq \ c_2^{-1} \ n^{-1} \ e^{-(S_n - L_n)} \ . \tag{4.71}$$

Thus $1 \leq 1 + U_n/V_n \leq 1 + c_1^{-1}$.

First, we treat the case $\mathbb{E}[Xe^{-X}] \leq 0$. As upper bound for the probability, we can estimate (as $L_n \leq 0$)

$$\mathbb{P}(Z_n = 1|\Pi) \quad \leq \quad c_1^{-2} e^{-S_n} e^{2L_n} \leq c_1^{-2} e^{-S_n} \tag{4.72}$$

and therefore

$$-\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}(Z_n = 1) \quad \geq \quad -\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{E}\left[e^{-S_n}\right] \quad = \quad -\log \mathbb{E}[e^{-X}] \ .$$

This proves the lower bound on ϱ.

For the lower bound of the probability (and the upper bound on ϱ) and the case $\mathbb{E}[Xe^{-X}] \geq 0$, using the estimates in (4.71) for some appropriate constant $0 < c < \infty$ yields

$$\mathbb{P}(Z_n = 1|\Pi) \quad \geq \quad c \ n^{-2} \ e^{-S_n} \ e^{2L_n} \ .$$

Thus,

$$
\begin{aligned}
-\liminf_{n\to\infty}\frac{1}{n}\log\mathbb{P}(Z_n=1) &\le -\liminf_{n\to\infty}\frac{1}{n}\log\mathbb{E}[c\;n^{-2}\;e^{-S_n}\;e^{2L_n}] \\
&= -\liminf_{n\to\infty}\frac{1}{n}\log\mathbb{E}[e^{-S_n}\;e^{2L_n}]\;. \qquad (4.73)
\end{aligned}
$$

From the preceding estimate, it follows that

$$
-\liminf_{n\to\infty}\frac{1}{n}\log\mathbb{P}(Z_n=1) \le -\liminf_{n\to\infty}\frac{1}{n}\log\mathbb{E}[e^{-S_n}\;e^{2L_n};L_n\le 0]\;.
$$

Next we change to the measure $\mathbf{P}$, for any bounded, measurable function $\phi:\mathbb{R}^n\to\mathbb{R}$ defined by (compare (1.6))

$$
\mathbf{E}[\phi(X_1,\dots,X_n)] := \mathbb{E}\big[\phi(X_1,\dots,X_n)e^{-S_n}\big]\mathbb{E}[e^{-X}]^{-n}\;.
$$

This yields

$$
\begin{aligned}
-\liminf_{n\to\infty}\frac{1}{n}\log\mathbb{P}(Z_n=1) &\le -\log\mathbb{E}\big[e^{-X}\big]-\liminf_{n\to\infty}\frac{1}{n}\log\mathbf{E}\big[e^{2L_n};L_n\ge 0\big] \\
&\le -\log\mathbb{E}\big[e^{-X}\big]-\liminf_{n\to\infty}\frac{1}{n}\log\mathbf{P}(L_n\ge 0)\;.
\end{aligned}
$$

Under $\mathbf{P}$, the associated random walk S has drift $\mathbf{E}[X]=\mathbb{E}[Xe^{-X}]\mathbb{E}[e^{-X}]^{-1}$. Thus, S has nonnegative drift under $\mathbf{P}$ if $\mathbf{E}[X]=\mathbb{E}[Xe^{-X}]\ge 0$ and the result follows for this case.

Secondly, in the case $\mathbb{E}[Xe^{-X}]<0$, by (4.73), it remains to prove

$$
-\frac{1}{n}\log\mathbb{E}[e^{-S_n+2L_n}] = \Lambda(0)\;.
$$

If $\mathbb{E}[Xe^{-X}]<0$ and $\mathbb{E}[X]>0$, there exists a $-1<\nu<0$ such that

$$
\mathbb{E}[Xe^{\nu X}] = 0\;.
$$

Next, we change to the measure $\widetilde{\mathbf{P}}$, for any bounded, measurable function $\phi:\mathbb{R}^n\to\mathbb{R}$ defined by (compare (1.6))

$$
\widetilde{\mathbf{E}}[\phi(X_1,\dots,X_n)] = \mathbb{E}[\phi(X_1,\dots,X_n)e^{\nu S_n}]\mathbb{E}[e^{\nu X}]^{-n}\;.
$$

Under $\widetilde{\mathbf{P}}$, S becomes a recurrent random walk. Thus we get

$$
-\frac{1}{n}\log\mathbb{E}[e^{-S_n+2L_n}] = -\log\mathbb{E}[e^{\nu X}]-\frac{1}{n}\log\widetilde{\mathbf{E}}[e^{(-\nu-1)S_n+2L_n}]\;. \qquad (4.74)
$$

Next, we will prove that the second term in the above sum vanishes. As $\nu>-1$, $e^{(-\nu-1)S_n+2L_n}$ is bounded from above by one. For the lower bound, we use the estimate

$$
\widetilde{\mathbf{E}}\big[e^{(-\nu-1)S_n+2L_n}\big] \ge \widetilde{\mathbf{E}}\big[e^{(-\nu-1)S_n};L_n\ge 0\big]\;.
$$

As $-\nu-1<0$, we may apply [ABKV10, Proposition 2.1] (see also Lemma 3.2.4), saying that the above expectation is of polynomial order. Thus $\lim_{n\to\infty}\frac{1}{n}\log\widetilde{\mathbf{E}}[e^{(-\nu-1)S_n+2L_n}]=0$ and by (4.71) and (4.74),

$$
\begin{aligned}
-\lim_{n\to\infty}\frac{1}{n}\log\mathbb{P}(Z_n=1) &= -\lim_{n\to\infty}\frac{1}{n}\log\mathbb{E}[e^{-S_n+2L_n}] \\
&= -\log\mathbb{E}[e^{\nu X}]\;.
\end{aligned}
$$

As $\Lambda(0)=-\log\mathbb{E}[e^{\nu X}]$, the lemma is proved. □

Lemma 4.4.4. *Let χ be defined as in (4.64). Then*

$$\chi(\theta) \quad = \quad \begin{cases} -\theta - \log \mathbb{E}\big[e^{-X}\big] & , \quad \text{if } \theta \le \theta^\star \\ \Lambda(\theta) & , \quad \text{if } \theta > \theta^\star \end{cases} \quad ,$$

where $\theta^\star = \mathbb{E}\big[Xe^{-X}\big]/\mathbb{E}\big[e^{-X}\big]$.

Proof. By Lemma 4.4.3, we only have to consider the case $\mathbb{E}[Xe^{-X}] \ge 0$, i.e. $\varrho = -\log\varphi(-1) \le \sup_{s\in\mathbb{R}}\{-\log\varphi(s)\} = \Lambda(0)$. By assumption, Λ is finite for $0 < \theta \le \mathbb{E}[X]$, thus differentiable. As $\Lambda(\mathbb{E}[X]) = 0$ and $\Lambda'(\mathbb{E}[X]) = 0$, there is a $0 \le \theta^\star \le \mathbb{E}[X]$ such that the tangent on Λ in $\theta^\star$ hits the ordinate in $-\log\varphi(-1)$. From representation (4.67) , χ follows this tangent for $\theta < \theta^\star$ and is identical to Λ for $\theta > \theta^\star$ (compare Section 4.3.6). It remains to prove $\theta^\star = \mathbb{E}[Xe^{-X}]/\mathbb{E}[e^{-X}]$ and $\Lambda'(\theta^\star) = -1$.

By definition, we have

$$\Lambda(\theta) \quad = \quad \sup_{s\in\mathbb{R}}\{\theta s - \log\varphi(s)\} \ .$$

If for some $\theta^\star$, the supremum is attained for $s = -1$, then

$$\Lambda(\theta^\star) \quad = \quad -\theta^\star - \log\varphi(-1) \ .$$

By theory of Legendre-transforms (see e.g. [dH00], the tangent on Λ in $\theta^\star$ is described by

$$\theta \quad \to \quad -\theta - \log\varphi(-1) \ ,$$

which proves the lemma. □

With a slight abuse of notation, we write below for the number of individuals $e^{\theta n}$ instead of the integer part $\lfloor e^{\theta n}\rfloor$.

Proof of the upper bound in Theorem 4.4.1. Let $0 < \theta < \mathbb{E}[X]$. Then

$$\mathbb{P}(1 \le Z_n \le e^{\theta n}) \quad \le \quad \mathbb{E}\big[\mathbb{P}(1 \le Z_n \le e^{\theta n}|\Pi); S_n > \theta n\big] + \mathbb{P}(S_n \le \theta n) \ . \tag{4.75}$$

We use the exact formula (4.61):

$$\begin{aligned} \mathbb{P}(1 \le Z_n \le e^{\theta n}|\Pi) \quad &= \quad (U_n+V_n)^{-1}\big(1 - (1+U_n/V_n)^{-e^{\theta n}}\big) \\ &\le \quad c_1^{-1}\Big(1 - \Big(1 + \frac{1}{\sum_{k=0}^{n-1} c_2 e^{S_n - S_k}}\Big)^{-e^{\theta n}}\Big) \\ &\le \quad c_1^{-1}\Big(1 - \Big(1 + \frac{1}{c_2\ n\ e^{S_n - L_n}}\Big)^{-e^{\theta n}}\Big) \\ &\le \quad \Big(c_1^{-1}c_2^{-1}\ n^{-1}\ e^{\theta n - (S_n - L_n)} + o\Big(n^{-1}\ e^{\theta n - (S_n - L_n)}\Big)\Big) \wedge 1 \ . \end{aligned}$$

Thus

$$\limsup_{n\to\infty} \frac{1}{n}\log\mathbb{E}\big[\mathbb{P}(1 \le Z_n \le e^{\theta n}|\Pi); S_n > \theta n\big] \quad \le \quad \limsup_{n\to\infty}\frac{1}{n}\log\mathbb{E}\big[e^{\theta n - S_n}; S_n > \theta n\big] \ .$$

Again, a change of measure is used. Recall the definition of $\mathbf{P}$ from Lemma 4.3,

$$\mathbf{E}[\phi(X_1,\ldots,X_n)] \quad = \quad \mathbb{E}\big[\phi(X_1,\ldots,X_n)e^{-S_n}\big]\mathbb{E}[e^{-X}]^{-n} \ .$$

Then

$$e^{\theta n}\ \mathbb{E}\big[e^{-S_n}; S_n > \theta n\big] \quad = \quad e^{\theta n}\ \mathbb{E}[e^{-X}]^n\ \mathbf{P}(S_n \ge \theta n) \ . \tag{4.76}$$

Now there are two cases. Under $\mathbf{P}$, either $\{S_n \ge \theta n\}$ is not a large deviation event (that is if $\mathbf{E}[X] = \mathbb{E}[Xe^{-X}]\mathbb{E}[e^{-X}]^{-1} \ge \theta$). Then

$$\lim_{n\to\infty}\frac{1}{n}\log\mathbf{P}(S_n \ge \theta n) \quad = \quad 0 \ .$$

Or it is a large deviation event (that is if $\mathbf{E}[X] = \mathbb{E}[Xe^{-X}]\mathbb{E}[e^{-X}]^{-1} < \theta$) and then

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbf{P}(S_n \geq \theta n) \quad =: \quad -\widetilde{\Lambda}(\theta) \ ,$$

where $\widetilde{\Lambda}$ is the rate function of S under $\mathbf{P}$. Now under $\mathbf{P}$, the moment generating function $\widetilde{\varphi}$ is

$$\widetilde{\varphi}(s) \quad := \quad \varphi(-1)^{-1}\mathbb{E}[e^{sX-X}] \ = \ \varphi(s-1)/\varphi(-1)$$

and the new rate function $\widetilde{\Lambda}$ becomes

$$\begin{aligned} \widetilde{\Lambda}(\theta) &= \sup_{s\in\mathbb{R}}\{\theta s - \log\widetilde{\varphi}(s)\} \\ &= \sup_{s\in\mathbb{R}}\{\theta s - \log\varphi(s-1)\} + \log\varphi(-1) \\ &= \Lambda(\theta) + \theta + \log\varphi(-1) \ . \end{aligned}$$

Then by (4.76), we have

$$\begin{aligned} &\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{E}\big[e^{\theta n - S_n}; S_n > \theta n\big] \\ &\leq \ - \inf_{t\in(0,1]} \begin{cases} -\theta - \log\varphi(-1) &, \ \text{if } \mathbb{E}[Xe^{-X}]\mathbb{E}[e^{-X}]^{-1} \geq \theta \\ \Lambda(\theta) &, \ \text{if } \mathbb{E}[Xe^{-X}]\mathbb{E}[e^{-X}]^{-1} < \theta \end{cases} \ . \end{aligned} \tag{4.77}$$

By standard arguments of large deviation theory (see Section 4.2) and (4.75),

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}(1 \leq Z_n \leq e^{\theta n}) \quad \leq \quad \begin{cases} \max\{\theta + \log\varphi(-1), -\Lambda(\theta)\} &, \ \text{if } \theta \leq \theta^\star \\ -\Lambda(\theta) &, \ \text{if } \theta > \theta^\star \end{cases} \ ,$$

where

$$\theta^\star \quad = \quad \mathbb{E}[Xe^{-X}]\mathbb{E}[e^{-X}]^{-1} \ .$$

Note that

$$\Lambda(\theta) \ = \ \sup_{s\in\mathbb{R}}\{s\theta - \log\varphi(s)\} \quad \geq \quad -\theta - \log\varphi(-1)$$

and thus

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}(1 \leq Z_n \leq e^{\theta n}) \quad \leq \quad \begin{cases} -\big(-\theta - \log\varphi(-1)\big) &, \ \text{if } \theta \leq \theta^\star \\ -\Lambda(\theta) &, \ \text{if } \theta > \theta^\star \end{cases} \ .$$

Together with Lemma 4.4.4, this proves the upper bound.

Proof of the lower bound in Theorem 4.4.1. For $\theta^\star < \theta < \mathbb{E}[X]$, the lower bound immediately follows from Lemma 4.4.2. As

$$\mathbb{P}(1 \leq Z_n \leq e^{\theta n}) \quad \geq \quad \sup_{t\in[0,1)} \Big\{\mathbb{P}(Z_{\lfloor tn\rfloor} = 1)\mathbb{P}(1 \leq Z_{\lfloor(1-t)n\rfloor} \leq e^{\theta n})\Big\}$$

and applying Lemma 4.4.2 yields

$$\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}(1 \leq Z_n \leq e^{\theta n}) \quad \geq \quad - \inf_{t\in[0,1)} \big\{t\varrho + (1-t)\Lambda(\theta/(1-t))\big\} \ .$$

This proves the theorem. □

4.5 The quenched approach

This section has been inspired by discussions with Nina Gantert[23]. The limit theorems of the previous section were discussed under the so-called annealed approach, i.e. the limits $n \to \infty$ were considered under the ***unconditioned*** measures. In this section, the asymptotics of the large deviation probabilities are studied ***conditioned on the environment***. As it turns out, the large deviation behavior is substantially different under the quenched approach. This is due to the fact that very improbable events might contribute to the expectation, but vanishe a.s. in the limit. For the results of this section, the bounds conditioned on the environment from the previous sections are used. Throughout this section, we assume $\mathbb{E}[|X|] < \infty$.

For the case of heavy-tailed offspring distributions, the following result holds:

Theorem 4.5.1. *If* $\limsup_{z\to\infty} \log \mathbb{P}(Z_1 > z|\Pi, Z_0 = 1)/\log z = -\beta$ *a.s. for some* $\beta \in (1,\infty)$ *and Assumption* $\mathcal{H}(\beta)$ *is fulfilled, then for every* $\theta \geq (\mathbf{E}[X] \vee 0)$*,*

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}(Z_n > e^{\theta n}|\Pi) \quad = \quad \begin{cases} -\beta(\theta - \mathbb{E}[X]) & , \quad \text{if} \quad \mathbb{E}[X] > 0 \\ -(\beta\theta - \mathbb{E}[X]) & , \quad \text{if} \quad \mathbb{E}[X] \leq 0 \end{cases} \qquad a.s.$$

This is comparable to the Galton-Watson case discussed in Section 4.3.1. In the supercritical case ($\mathbb{E}[X] > 0$), essentially large deviations are attained by a jump in the first generation of size $(\theta - \mathbb{E}[X])$. After that, the process grows according to its expectation (see Figure 4.9).

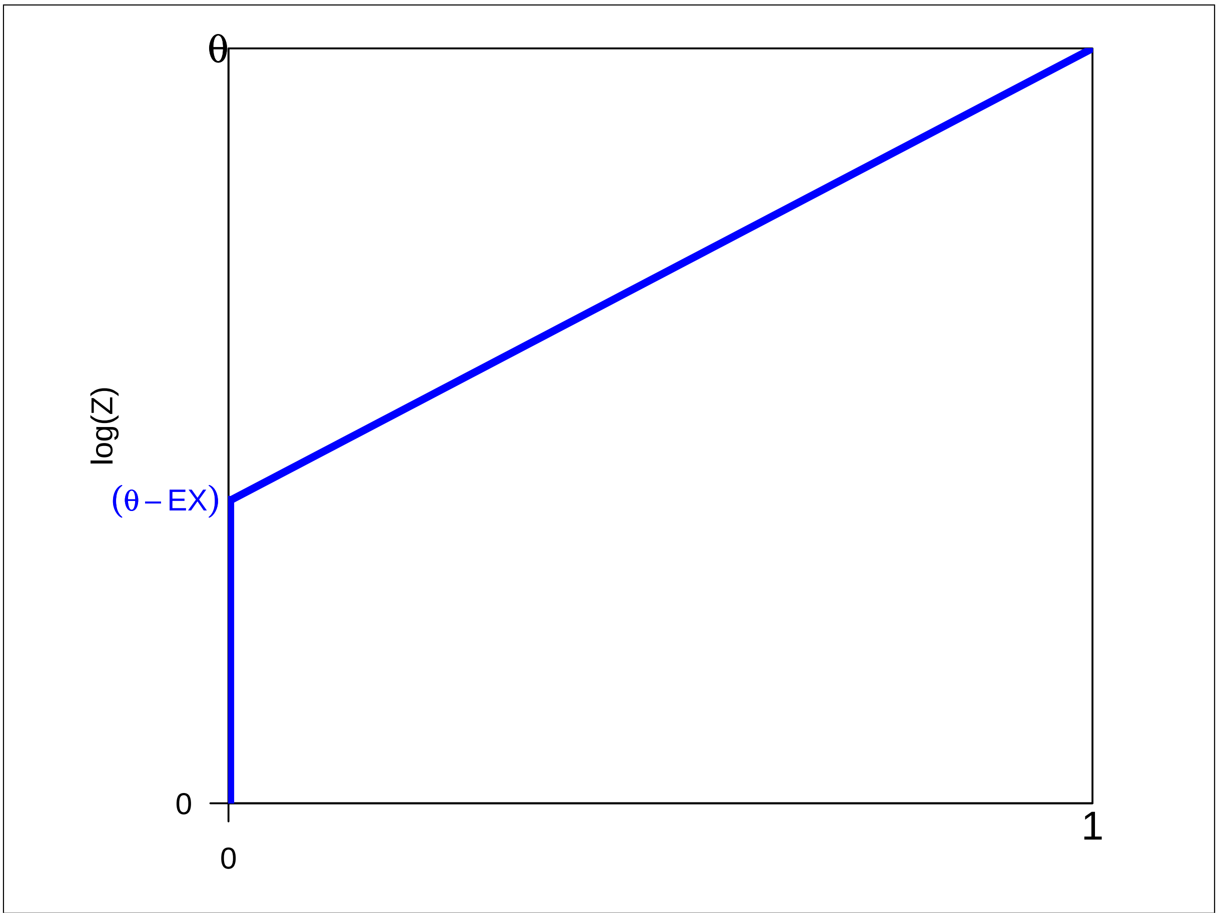


Figure 4.9: Typical path for large deviations of a supercritical BPRE (quenched approach, scaled by n).

[23]Universität Münster

In the case of a critical or subcritical BPRE ($\mathbb{E}[X] = 0$ or $\mathbb{E}[X] < 0$),

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}(Z_n > e^{\theta n}) \quad = \quad -(\gamma + \beta\theta) \quad \text{a.s.}$$

with $\gamma := -\mathbb{E}[X]$. As we will see later, under Assumption $\mathcal{H}(\beta)$ for critical and subcritical BPREs,

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}(Z_n > 0|\Pi) \quad = \quad -\gamma \ .$$

Again, compare this with the Galton-Watson case discussed in Section 4.3.1. By the strong law of large numbers, $S_n = n\mathbb{E}[X] + o(n)$ a.s. and in the limit, the associated random walk behaves deterministically and is linear with slope $\mathbb{E}[X]$. The process just survives until the end (having exponentially small probability of order $e^{\mathbb{E}[X]n}$) and then has a jump of size $e^{\theta n}$, which also has exponentially small probability of order $e^{-\beta\theta n}$.

If the offspring distributions have geometrically bounded tails, the probability of attaining exceptionally small values is of lower order than exponential. This is described by our next theorem.

Theorem 4.5.2. *If $\mathbb{P}(Z_1 > e^{\theta}|\Pi) > 0$ a.s., under Assumption 4.2, for every $\theta \geq (\mathbf{E}[X] \vee 0)$, the following holds*

$$\lim_{n\to\infty} \frac{1}{n} \log(-\log \mathbb{P}(Z_n > e^{\theta n}|\Pi)) \quad = \quad \begin{cases} \theta - \mathbb{E}[X] & , \ \text{if} \quad \mathbb{E}[X] > 0 \\ \theta & , \ \text{if} \quad \mathbb{E}[X] \leq 0 \end{cases} \qquad \text{a.s.}$$

Thus, for a supercritical BPRE,

$$\mathbb{P}\Big(\mathbb{P}(Z_n > e^{\theta n}|\Pi) = e^{-e^{(\theta-\mathbb{E}[X])n+o(n)}}\Big) \overset{n\to\infty}{\longrightarrow} 1$$

and for a critical or subcritical BPRE,

$$\mathbb{P}\Big(\mathbb{P}(Z_n > e^{\theta n}|\Pi) = e^{-e^{\theta n+o(n)}}\Big) \overset{n\to\infty}{\longrightarrow} 1$$

Note that if $\mathbb{P}(Z_1 > e^{\theta}|\Pi) = 0$ a.s., then $\mathbb{P}(Z_n > e^{\theta n}|\Pi) = 0$.

As we will see in the proof of this Theorem, the result can be interpreted as follows. Conditioned on the environment, the associated random walk behaves in the limit deterministically and is linear with slope $\mathbb{E}[X]$. Thus, the only strategy to attain exceptionally large values is to affect the branching mechanism. Essentially, large deviations are attained by $\lceil e^{\theta k}\rceil$-many individuals in each generation k having at least e^{θ}-many offsprings ($0 \leq k \leq n$). As, conditioned on the environment, the branching is independent, the cost for this strategy in generation k is $(\mathbb{P}(Z_k > e^{\theta}|\Pi))^{\lceil \exp(\theta k)\rceil}$. Multiplying the probabilities for each generation, we end up with a 'double-exponential' cost as claimed in Theorem 4.5.2. Note that in the limit, the exact value of $\mathbb{P}(Z_1 > e^{\theta}|\Pi) > 0$ is not of importance.

4.5.1 Proof of Theorems 4.5.1 and 4.5.2

For our proofs, the following result for the a.s.-limit of a reflected random walk is required.

Lemma 4.5.3. *Let $(S_n)_{n\geq 0}$ be a random walk with increments distributed like X. If $\mathbb{E}[|X|] < \infty$, then*

$$\lim_{n\to\infty} \frac{1}{n}(S_n - L_n) \quad = \quad (\mathbb{E}[X] \vee 0) \quad \text{a.s.}$$

Proof. It suffices to prove that $L_n/n \to \mathbb{E}[X] \wedge 0$ a.s. Rewriting the minimum L_n yields for any $1 \leq m \leq n$,

$$\begin{aligned} \frac{L_n}{n} &= \min\Big\{\frac{1}{n}\frac{S_1}{1}, \frac{2}{n}\frac{S_2}{2}, \cdots, \frac{n}{n}\frac{S_n}{n}\Big\} \\ &= \min\Big\{\min_{k=1,\dots,m}\Big\{\frac{k}{n}\frac{S_k}{k}\Big\}, \min_{k=m+1,\dots,n}\Big\{\frac{k}{n}\frac{S_k}{k}\Big\}\Big\} \ . \end{aligned} \tag{4.78}$$

By the strong law of large numbers, $S_n/n \to \mathbb{E}[X]$ a.s. (see e.g. [Kal01, Theorem 4.23]) and thus, for any $\epsilon > 0$, there is a $N < \infty$ a.s. such that $|S_n/n - \mathbb{E}[X]| < \epsilon$ for all $n \geq N$. By (4.78), for all $n \geq N$,

$$\frac{L_n}{n} \quad \geq \quad \min\left\{ \min_{k=1,\dots,N} \left\{\frac{k}{n}\frac{S_k}{k}\right\}, \min_{k=N+1,\dots,n} \left\{\frac{k}{n}\mathbb{E}[X]\right\}\right\} - \epsilon \ .$$

Taking the limit $n \to \infty$ yields

$$\frac{L_n}{n} \quad \geq \quad 0 \wedge \mathbb{E}[X] - \epsilon \quad \text{a.s.}$$

Letting $\epsilon \to 0$ ensues the lower bound. The upper bound follows trivially by $L_n/n \leq \min\{S_1/n, S_n/n\}$. □

For the proof of Theorem 4.5.1, the following lemma is required. It describes that Z_n is –with a high probability– as large as its expectation, if L_n is large.

Lemma 4.5.4. *Under Assumption $\mathcal{H}(\beta)$ and for $\beta \in (1, 2]$ and $r \in (0, 1)$, there is a constant $c > 0$ such that*

$$\mathbb{P}\big(Z_n \geq r\mathbb{E}[Z_n|\Pi]\big|\Pi\big) \quad \geq \quad c\ (1-r)^{\beta/(\beta-1)} n^{-\lceil\beta\rceil/(\beta-1)} e^{L_n}\ \Upsilon(e^{-L_n}n^2)^{-1/(\beta-1)} \ .$$

For $\beta > 2$,

$$\mathbb{P}\big(Z_n > r\mathbb{E}[Z_n|\Pi|]\big|\Pi\big) \quad \geq \quad c\ (1-r)^2\ e^{L_n} \quad \textit{a.s.}$$

Proof. The second statement is a consequence of (4.38). If $\mathbb{E}[X] > 0$ and for $\beta > 2$, by (4.38),

$$\mathbb{E}[Z_n(Z_n - 1)] \quad = \quad f_{0,n}^{(2)}(1) \ \leq \ c\ e^{S_n}e^{S_n - L_n} \quad \text{a.s.}$$

The inequality due to Paley and Zygmund (see e.g. (4.21)) yields the result.

For $\beta \in (1, 2]$, a more general form of (4.21) is used, which is proved in the appendix (see Lemma A.1.1). For $r \in (0, 1)$ and $p, q > 0$, $p + q = 1$:

$$\mathbb{P}(Y \geq r\mathbb{E}[Y])^q \quad \geq \quad \frac{(1-r)\mathbb{E}[Y]}{\mathbb{E}[Y^{1/p}]^p} \ .$$

For any $0 < \iota < \beta - 1$, let $1/p = 1 + \iota$. As

$$\mathbb{E}[Y^{1+\iota}] \quad = \quad \frac{1}{1+\iota}\int_0^\infty y^\iota \mathbb{P}(Y > y)dy \ ,$$

we can estimate $\mathbb{E}\big[Z_n^{1+\iota}\big|\Pi\big]$ using Theorem 4.3.4:

$$\mathbb{E}\big[Z_n^{1+\iota}\big|\Pi\big] \quad \leq \quad c\ n^{\lceil\beta\rceil}\ e^{S_n}\ e^{(\beta-1)(S_n-L_n)} \sum_{y=1}^\infty y^{\iota-\beta}\Upsilon(n^2e^{-L_n}y) \ .$$

By properties of slowly varying sequences (see appendix) and as $\iota - \beta < -1$, the integral is finite and can be bounded by $c\Upsilon(n^2e^{-L_n})$ for some constant c. As $p/q = 1/\iota$, $1/q = 1 + 1/\iota$ and thus, for some constant c,

$$\mathbb{P}\big(Z_n \geq r\mathbb{E}[Z_n|\Pi]\big|\Pi\big) \quad \geq \quad c\ n^{-\lceil\beta\rceil/\iota}\ (1-r)^{1+1/\iota}e^{S_n}e^{-(S_n-L_n)(\beta-1)/\iota}\ \Upsilon(e^{-L_n}n^2)^{-1/\iota} \ .$$

Taking the limit $\iota \to \beta - 1$ yields the result. □

Proof of Theorem 4.5.1. Let $\theta > \mathbb{E}[X] \vee 0$. First, assume $\beta \in (1, 2]$. Theorem 4.3.4 yields

$$\frac{1}{n}\log\mathbb{P}(Z_n > e^{\theta n}|\Pi) \quad \leq \quad \frac{S_n + (\beta-1)(S_n - L_n)}{n} - \beta\theta + o(1) \quad \text{a.s.}$$

and by Lemma 4.5.3 and the strong law of large numbers,

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}(Z_n > e^{\theta n}|\Pi) \quad \leq \quad -\beta\theta + \mathbb{E}[X] + (\beta-1)(\mathbb{E}[X] \vee 0) \quad \text{a.s.}$$

If $\beta > 2$, Theorem 4.3.6 can be used:

$$\begin{aligned} &\frac{1}{n} \log \mathbb{P}(Z_n > e^{\theta n}|\Pi) \\ &\quad \leq \max\Big\{ \frac{S_n + (\beta-1)(S_n - L_n)}{n} - \beta\theta; \frac{S_n + \lceil\beta\rceil(S_n - L_n)}{n} - (\lceil\beta\rceil + 1)\theta \Big\} + o(1) \quad \text{a.s.} \end{aligned}$$

Letting $n \to \infty$, by the strong law of large numbers and as $\theta > \mathbb{E}[X]$

$$\begin{aligned} &\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}(Z_n > e^{\theta n}|\Pi) \\ &\quad \leq \max\big\{ -\beta\theta + \mathbb{E}[X] + (\beta-1)(\mathbb{E}[X] \vee 0); -(\lceil\beta\rceil + 1)\theta + \mathbb{E}[X] + \lceil\beta\rceil(\mathbb{E}[X] \vee 0)\big\} \quad \text{a.s.} \\ &\quad = -\beta\theta + \mathbb{E}[X] + (\beta-1)(\mathbb{E}[X] \vee 0) \quad \text{a.s.} \end{aligned}$$

This proves the upper bound in Theorem 4.5.1.

For the lower bound, first consider the case $\mathbb{E}[X] \leq 0$. By (4.29) and Lemma 4.5.3

$$\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}(Z_{n-1} > 0|\Pi) \ \geq \ \liminf_{n\to\infty} \frac{L_n}{n} \ = \ \mathbb{E}[X] \ .$$

The estimate

$$\mathbb{P}(Z_n > e^{\theta n}|\Pi) \ \geq \ \mathbb{P}(Z_{n-1} > 0|\Pi)\mathbb{P}(Z_n > e^{\theta n}|\Pi, Z_{n-1} = 1)$$

together with the assumption yields the claim for $\mathbb{E}[X] \leq 0$.

For the case $\mathbb{E}[X] > 0$, we follow the same ideas as in Section 4.3.2 and exhibit the optimal strategy which starts with a jump. Here, this means that in the first generation, an individual has $e^{(\theta-\mathbb{E}[X])n}$-many children. By assumption, the probability for this event is asymptotically

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}\big(Z_1 > \exp((\theta - \mathbb{E}[X])n)\big) = -\beta(\theta - \mathbb{E}[X]) \ . \tag{4.79}$$

As

$$\mathbb{P}(Z_n > e^{\theta n}|\Pi) \ \geq \ \mathbb{P}\big(Z_1 > e^{(\theta-\mathbb{E}[X])n}\big|\Pi\big)\mathbb{P}\big(Z_n > e^{\theta n}\big|\Pi, Z_1 = e^{(\theta-\mathbb{E}[X])n}\big) \ , \tag{4.80}$$

it remains to prove that

$$\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}\big(Z_n > e^{\theta n}|\Pi, Z_1 = e^{(\theta-\mathbb{E}[X])n}\big) \ = \ 0 \ .$$

Lemma 4.5.4 ensues for any $r \in (0,1)$ and $\beta \in (1,2]$

$$p_n \ := \ \mathbb{P}\big(Z_n > r\mathbb{E}[Z_n|\Pi]\big|\Pi\big) \quad \geq \quad c\,(1-r)^{\beta/(\beta-1)} n^{-\lceil\beta\rceil/(\beta-1)} e^{L_n} \ \Upsilon(e^{-L_n}n^2)^{-1/(\beta-1)} \quad \text{a.s.}$$

By the same arguments as in Section 4.3.2,

$$\mathbb{P}\Big(N_n \geq \frac{\mathbb{E}[N_n \mid \Pi]}{Np_n} \ \Big|\ \Pi, Z_s = e^{sn}\Big) \quad \geq \quad \frac{\Big[1 - 1 \wedge \frac{1}{Np_n}\Big]^2}{1 + \frac{e^{-sn}}{p_n}} \ , \tag{4.81}$$

where $N \geq 1$ and N_n is number of the e^{sn}-many subtrees (i) with $Z_n^{(i)} > Ne^{\theta n}$. As for $\mathbb{E}[X] > 0$, L_n is of constant order, we conclude from (4.81)

$$\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}\big(Z_n > e^{\theta n}|\Pi, Z_1 = e^{(\theta-\mathbb{E}[X])n}\big) \ = \ 0 \ .$$

The same arguments can be applied for $\beta > 2$ (by using Lemma 4.5.4). Thus, by (4.80) and (4.79),

$$\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}(Z_n > e^{\theta n}) \ \geq\ -\beta(\theta - \mathbb{E}[X]) \ ,$$

which proves the Theorem. □

Proof of Theorem 4.5.2. Let $\theta > \mathbb{E}[X] \vee 0$. For the lower bound, we use Theorem 4.2.4 which yields

$$\frac{1}{n} \log\big(- \log \mathbb{P}(Z_n > e^{\theta n} | \Pi)\big) \quad \geq \quad n^{-1} \log \Big(\frac{U_n}{V_n}\Big) + \theta + o(n^{-1}) \ .$$

Recall $U_n = e^{-S_n}$ and, as we have seen in Section 4.2 under Assumption 4.2, $V_n \leq (n+1)e^{-L_n}$. Together with Lemma 4.5.3, this yields

$$\liminf_{n\to\infty} \frac{1}{n} \log(- \log \mathbb{P}(Z_n > e^{\theta n} | \Pi)) \quad \geq \quad \theta - (\mathbb{E}[X] \vee 0) \quad \text{a.s.}$$

For the upper bound, let $p > 0$ be such that $\mathbb{P}(Z_1 > e^{\theta} | \Pi) > p$ a.s. First let $\mathbb{E}[X] \leq 0$. Then the probability of having at least $e^{\theta n}$ many individuals in generation n may be estimated by

$$\mathbb{P}(Z_n > e^{\theta n} | \Pi) \quad \geq \quad \prod_{k=1}^{n} \mathbb{P}(Z_k > e^{\theta} | \Pi, Z_{k-1} = 1)^{e^{\theta k}} \ \geq \ \prod_{k=1}^{n} p^{e^{\theta k}} \ .$$

Thus

$$\begin{aligned}
\frac{1}{n} \log(- \log \mathbb{P}(Z_n > e^{\theta n} | \Pi)) \quad &\leq \quad \frac{1}{n} \log \big(\sum_{k=1}^{n} e^{\theta k}(- \log p)\big) \\
&\leq \quad \frac{1}{n} \log \big(n e^{\theta n}\big) \\
&\leq \quad \theta + \frac{\log n}{n} \ .
\end{aligned}$$

Now taking the limes superior yields the result.

If $\mathbb{E}[X] > 0$, we estimate

$$\begin{aligned}
\mathbb{P}(Z_n > e^{\theta n} | \Pi) \ &\geq \ \mathbb{P}\big(Z_1 > e^{\theta} \big| \Pi\big) \cdot \mathbb{P}\big(Z_n > e^{\mathbb{E}[X] n} \big| \Pi, Z_1 = e^{\theta - \mathbb{E}[X]}\big) \cdot \mathbb{P}\big(Z_n > e^{(\theta - \mathbb{E}[X]) n} \big| \Pi, Z_1 = e^{\mathbb{E}[X]}\big) \\
&=: p_{1n} \cdot p_{2n} \cdot p_{3n} \ .
\end{aligned}$$

Then standard arguments from large deviation theory (see [dH00]) yield

$$\begin{aligned}
\limsup_{n\to\infty} \frac{1}{n} \log\big(- \log \mathbb{P}(Z_n > e^{\theta n} | \Pi)\big) \ &\leq \ \limsup_{n\to\infty} \frac{1}{n} \log\big(- \log p_{1n} - \log p_{2n} - \log p_{3n}\big) \\
&= \ \max_{i=1,2,3} \Big\{ \limsup_{n\to\infty} \frac{1}{n} \log\big(- \log p_{in}\big) \Big\} \ .
\end{aligned}$$

p_{1n} is of constant order. Using Lemma 4.5.4 for $r \to 1$ and as L_n is of constant order for a random walk with positive drift, $\mathbb{E}[X] > 0$, also p_{3n} is of constant order. Using the calculation above,

$$\begin{aligned}
\limsup_{n\to\infty} \frac{1}{n} \log\big(- \log p_{2n}\big) \ &= \ \limsup_{n\to\infty} \frac{1}{n} \log\big(- \log \mathbb{P}(Z_n > e^{(\theta - \mathbb{E}[X]) n} | \Pi)\big) \\
&= \ \theta - \mathbb{E}[X] \ ,
\end{aligned}$$

which proves the theorem. □

Chapter 5

Simulation of a conditioned BPRE

In this section, a routine to simulate a BPRE conditioned on survival (realized in the statistical program R)[24] is described. First, the environment is (randomly) created and then a Galton-Watson tree in a given environment is simulated. For modelling a conditioned Galton-Watson process in a varying environment, Geiger's construction (see [Gei99, Lemma 2.1]) is used.

Later, only geometric offspring distributions are considered. Then the explicit calculation of the distribution of Z_n, conditioned on the environment, is feasible.

5.1 Geiger's construction for Galton-Watson processes in varying environment

In this section, it is assumed that the environment is given. Let us briefly explain the Geiger construction (see [Gei99, Lemma 2.1]). It says that a Galton-Watson process, conditioned on survival until generation n, can be built up in the following way:

We follow the 'line of descent 'of the leftmost individual having a descendant in generation n. Let $\mathbb{L}_k$ be the number of the leftmost individual in generation k, having an ancestor in generation n. To the left of $\mathbb{L}_k$, independent subtrees conditioned on extinction in generation n are growing. To the right of $\mathbb{L}_k$, independent unconditioned trees are evolving. This is illustrated in Figure 5.1. The joint distribution of $\mathbb{L}_k$ and the number of offsprings in generation k is known. Let us explain this in detail:

Define for $0 \leq k < n$

$$\begin{aligned} p_{k,n} &:= \mathbb{P}(Z_n > 0 | Z_k = 1, \Pi) \\ q_{k,n} &:= \mathbb{P}(Z_n = 0 | Z_k = 1, \Pi) \;=\; 1 - p_{k,n} \;. \end{aligned}$$

Then Geiger's construction goes as follows. For convenience, let $Z_0 = 1$. In the first generation, the joint distribution of $(\mathbb{L}_1, Z_1)$, conditioned on $\{Z_n > 0\}$, is given by

$$\begin{aligned} \mathbb{P}(\mathbb{L}_1 = j, Z_1 = i | Z_n > 0, \Pi) &= Q_1(i) \frac{\mathbb{P}(Z_n > 0 | Z_1 = 1, \Pi) \mathbb{P}(Z_n = 0 | Z_1 = j-1, \Pi)}{\mathbb{P}(Z_n > 0 | \Pi)} \\ &= Q_1(i) \frac{p_{1,n} q_{1,n}^{j-1}}{p_{0,n}} \quad , \; (1 \leq j \leq i) \;. \end{aligned}$$

More generally, for every $0 \leq k < n$,

$$\mathbb{P}(\mathbb{L}_k = j, Z_k = i | Z_n > 0, Z_{k-1} = 1, \Pi) \;=\; Q_k(i) \frac{p_{k,n} q_{k,n}^{j-1}}{p_{k-1,n}} \quad , \; (0 \leq j \leq i) \;. \tag{5.1}$$

The individuals to the left of $\mathbb{L}_k$ give rise to independent subtrees, conditioned on extinction before generation n, the individuals to the right found independent unconditioned subtrees. Denote by $Z_k^{(c)}$ the

[24] see the *R Projekt for Statistical Computing*, www.r-project.org

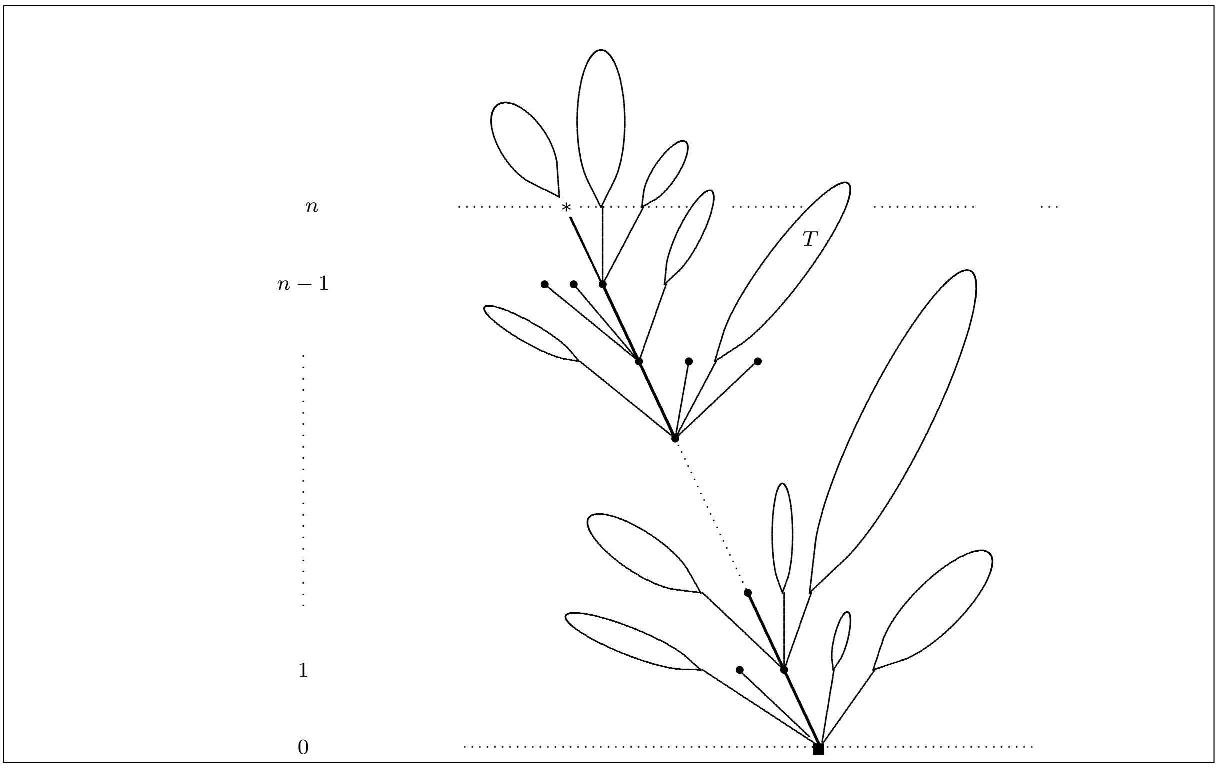


Figure 5.1: Illustration of Geiger's construction of a conditioned Galton-Watson process (taken from [BGK05]).

number of indivduals in generation k in the conditioned subtrees, and by $Z_k^{(u)}$ the number of individuals in generation k in the unconditioned subtrees. By Y_k, the number of children of the 'line of descent' in generation k is denoted. Thus the conditioned Galton-Watson tree in varying environment can be modelled in the following way:

1. Choose randomly $\mathbb{L}_k$ and Y_k with distribution (5.1).

2. Each of the $Z_{k-1}^{(c)}$-many individuals from the conditioned subtrees independently gives birth to a random number of children, with distribution

$$d(j) \ := \ Q_k(j) \cdot \frac{q_{k,n}^j}{q_{k-1,n}} \ . \tag{5.2}$$

   Each of the $Z_k^{(u)}$-many individuals has an independent number of children, with distribution Q_k.

3. The $(\mathbb{L}_k - 1)$-many individuals are added to the ***conditioned*** subtrees, the $(Y_k - \mathbb{L}_k)$-many individuals are added to the ***unconditioned*** subtrees.

4. Start afresh with step 1 in generation $k + 1$.

5.2 Geometric offspring distributions

Assume that the offspring distributions are geometric. In this case (see e.g. [Koz06]),

$$\mathbb{P}(Z_n > 0|\Pi, Z_0 = 1) \;=\; \frac{1}{e^{-S_n} + \sum_{j=0}^{n-1} e^{-S_j}} \;=\; \frac{1}{\sum_{j=0}^{n} e^{-S_j}} \;.$$

More generally, as the branching is independent conditioned on Π,

$$\begin{aligned}\mathbb{P}(Z_n > 0|\Pi, Z_i = k) \;&=\; 1 - \big(1 - \mathbb{P}(Z_n = 0|\Pi, Z_i = 1)\big)^k \\ &=\; 1 - \Big(1 - \frac{1}{\sum_{j=i}^{n} e^{-(S_j - S_i)}}\Big)^k \;.\end{aligned}$$

Thus, the probabilities $q_{k,n}$ and $p_{k,n}$ can be easily calculated:

$$\begin{aligned}p_{k,n} \;&=\; \frac{1}{\sum_{j=k}^{n} e^{-(S_j - S_k)}} \\ q_{k,n} \;&=\; 1 - \frac{1}{\sum_{j=k}^{n} e^{-(S_j - S_k)}} \;.\end{aligned}$$

Another advantage of the geometric offspring distributions is that the conditioned distribution (5.2) is still geometric. Let $p_k = p(Q_k)$ be the parameter of a geometric distribution in generation k. Then

$$d(j) \;=\; p_k \cdot (1 - p_k)^j \cdot \frac{q_{k,n}^j}{q_{k,n}} \;=\; \frac{p_k}{q_{k-1,n}} \big((1 - p_k) q_{k,n}\big)^j$$

defines a geometric distribution with parameter $1-(1-p_k)q_{k,n}$. Thus, the total offspring number of all conditioned subtrees in generation k is negative binomial distributed with parameters $(Z_{k-1}^{(c)}, 1-(1-p_k)q_{k,n})$. The total offspring number of all unconditioned subtrees in generation k is also negative binomial distributed with parameters $(Z_{k-1}^{(u)}, p_k)$.

5.3 Conditioned BPREs

In this section, the purpose is to simulate an intermediately subcritical BPRE, conditioned on $\{Z_n > 0\}$. As a direct simulation is hardly feasible (the event $\{Z_n > 0\}$ has exponentially small probability), an approximation is made, using Theorem 3.1.3:

First an oscillating random walk S, conditioned to have its minimum at the end, is picked randomly. Due to Theorem 3.1.3, the associated random walk conditioned on survival of the process asymptotically looks like a process conditioned on having its minimum at the end. Thus this approach is asymptotically correct.

In a second step, a conditioned Galton-Watson process is simulated in the environment given by the conditioned random walk, following the scheme described in the preceding section. The parameter of the geometric offspring distribution p_k (of the unconditioned process) in generation k is then given by the corresponding increment of the associated random walk:

$$p_k \;=\; \frac{1}{e^{S_k - S_{k-1}} + 1} \;.$$

For simplicity, only simple, symmetric random walks are considered, with just two different states of the environment. In this case, the simulation of the conditioned random walk is simple. First, a random walk conditioned on $\{M_n < 0\}$ is created randomly. Using the usual Doob h-transform (see [RW00, Chapter III.28]) for the conditioning, the transition probabilities of the random walk conditioned on $\{M_n < 0\}$ become

$$\begin{aligned}\mathbb{P}(S_1 = -1|M_n < 0) =&\; 1 \\ \mathbb{P}(S_k = S_{k-1} + 1|M_n < 0) =&\; \frac{S_{k-1} + 1}{2S_{k-1}} \qquad , \; 2 \le k \le n \;.\end{aligned}$$

By going to the dual random walk $(\tilde{S}_k)_{0\le k\le n}$, defined by

$$\tilde{S}_k \;:=\; S_n - S_{n-k} \qquad , \; 0 \le k \le n \;,$$

one gets the desired process, conditioned on having its minimum at the end.

5.4 Some results of the simulations

Several aspects of conditioned BPREs may be illustrated by the simulation scheme described in the preceding sections. As the corresponding limit theorems for other critical and subcritical cases are known (see Chapter 2), a similiar scheme can be used for those cases. Here, we only focus on the intermediately subcritical case. In Figure 5.2, a random walk conditioned on having its minimum at the end is displayed. It serves as environment for the intermediately BPRE, conditioned on $\{Z_{1000} > 0\}$. The logarithm of the number of individuals of the corresponding conditioned BPRE is displayed in Figure 5.3.

In the proof of Theorem 3.1.2, the fact was used that the number of individuals at the strict descending ladder points, $(\gamma_i)_{i\in\mathbb{N}}$, is bounded. Figure 5.4 displays the number of individuals at times, when the associated random walk has a strictly descending ladder point. As we can see, there are only very few individuals. In the supercritical periods between those *bottlenecks*, the number of individuals may attain very large values (10^{10}).

One might guess that the number of individuals may depend on the length of the excursion between two strictly descending ladder points. After a long excursion, there may be a tendency to have more individuals. Figure 5.6 displays the number of individuals in the bottlenecks (*blue*), averaged over 1000 simulations and the excursions lengths (*red*) between the strictly descending ladder points. This simulation has been conducted in the environment displayed in Figure 5.5.

Finally, the open problem in Chapter 3.3 can be illustrated. Let $(S^r_n)_{n\in\mathbb{N}}$ be the reflected random walk, defined by

$$S^r_n \;:=\; S_n - L_n \;. \tag{5.3}$$

During the excursions between the strictly descending ladder points, it is expected that Z/e^{S^r} converges to some finite random variable W. This convergence is illustrated in Figure 5.7, where the corresponding environment is displayed in Figure 5.2. It is clearly visible that Z/e^{S^r} converges to some random variable if the excursion is long enough.

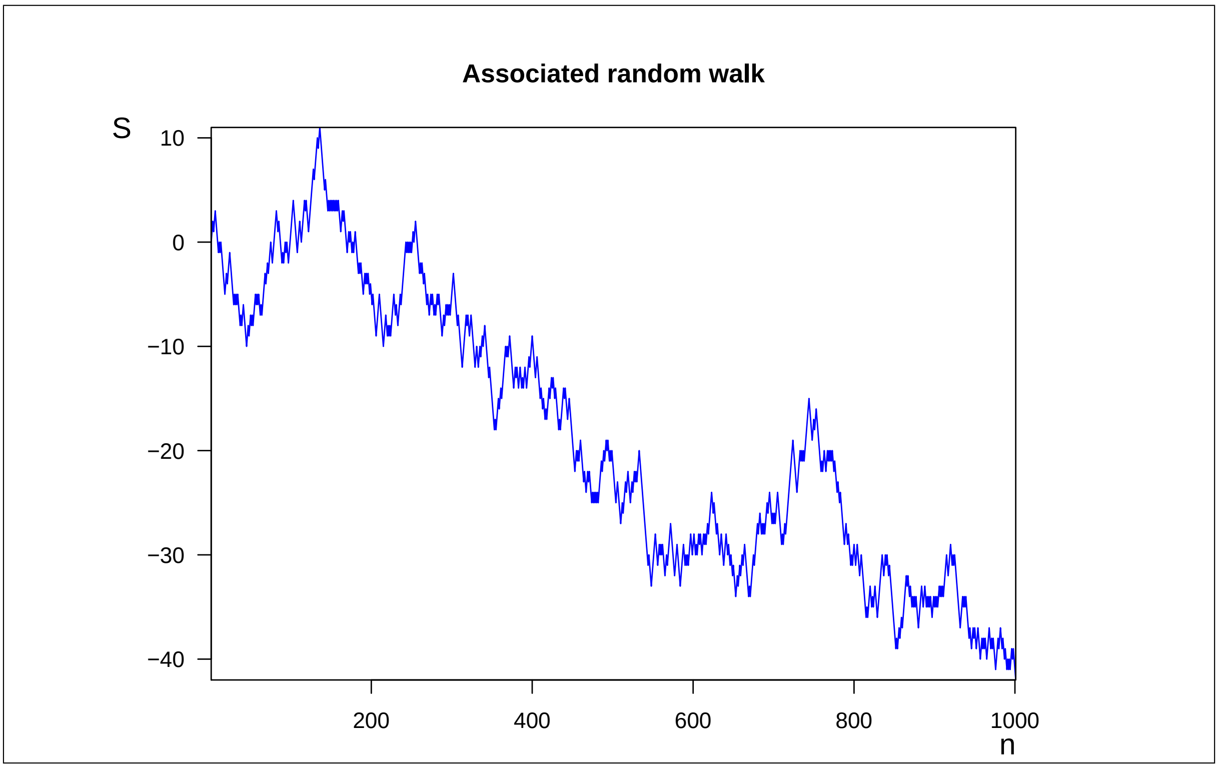


Figure 5.2: Example: Associated random walk.

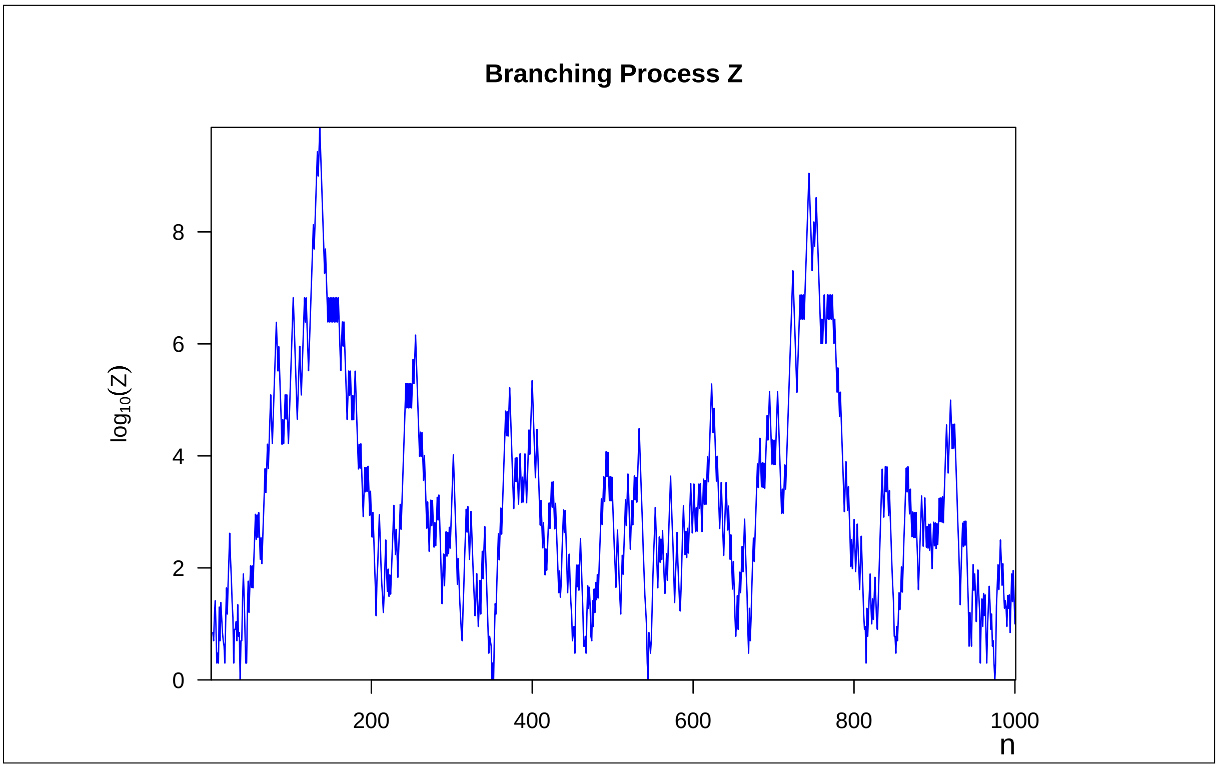


Figure 5.3: Example: $\log_{10}(Z)$ of the BPRE Z, conditioned on $\{Z_{1000} > 0\}$, in the environment displayed in Figure 5.2.

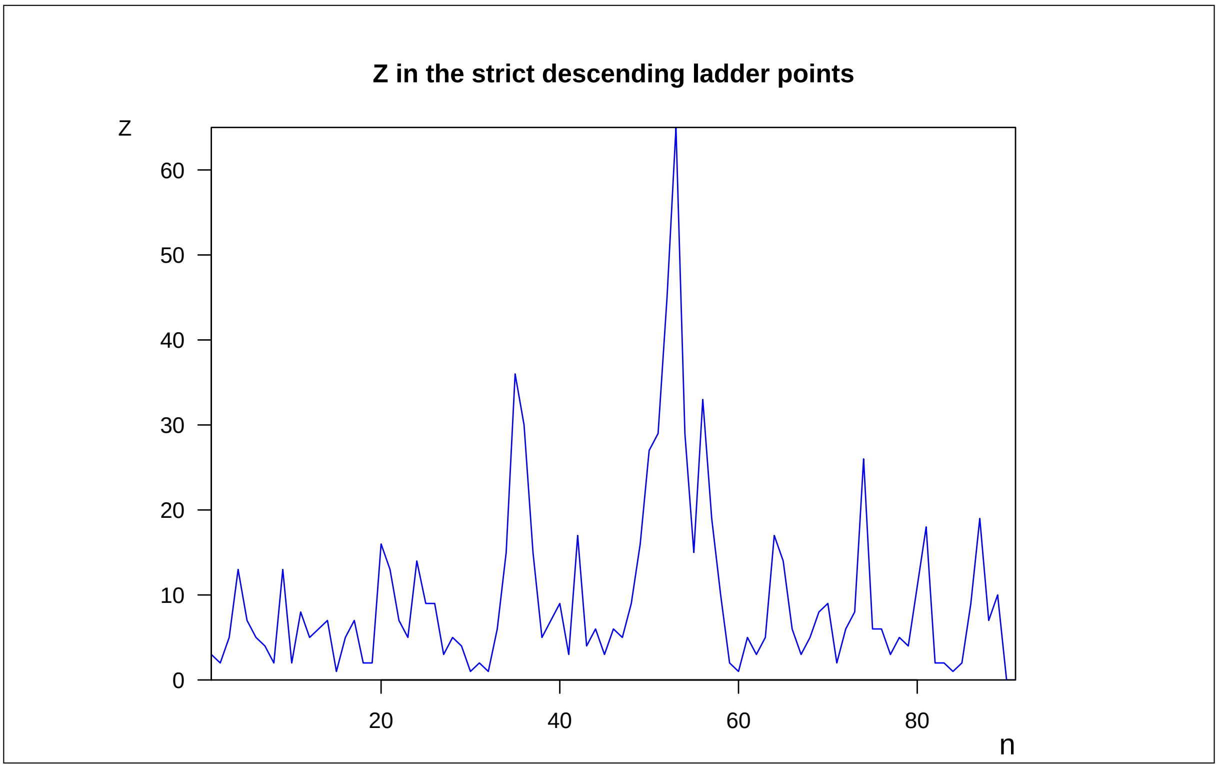


Figure 5.4: Example: Z, conditioned on $\{Z_{1000} > 0\}$, at the strictly descending ladder times of the associated random walk in Figure 5.2.

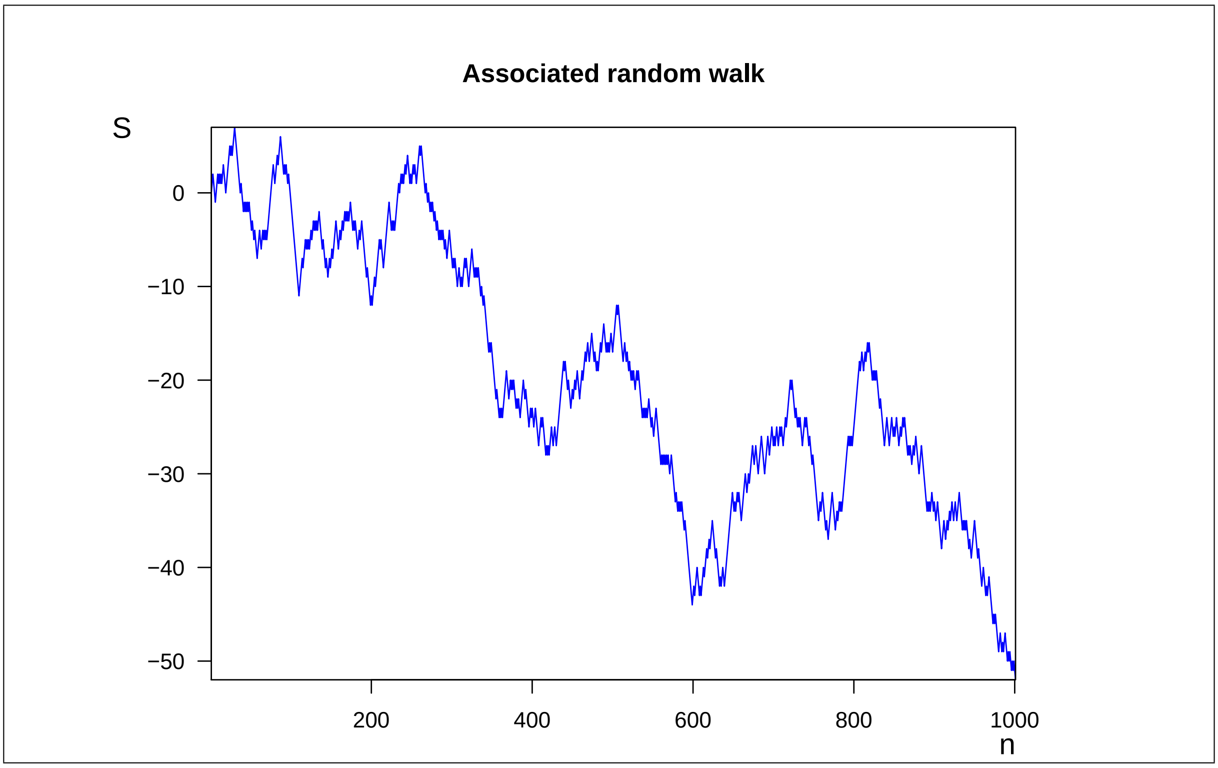


Figure 5.5: Example: Associated random walk, conditioned on having its minimum at the end.

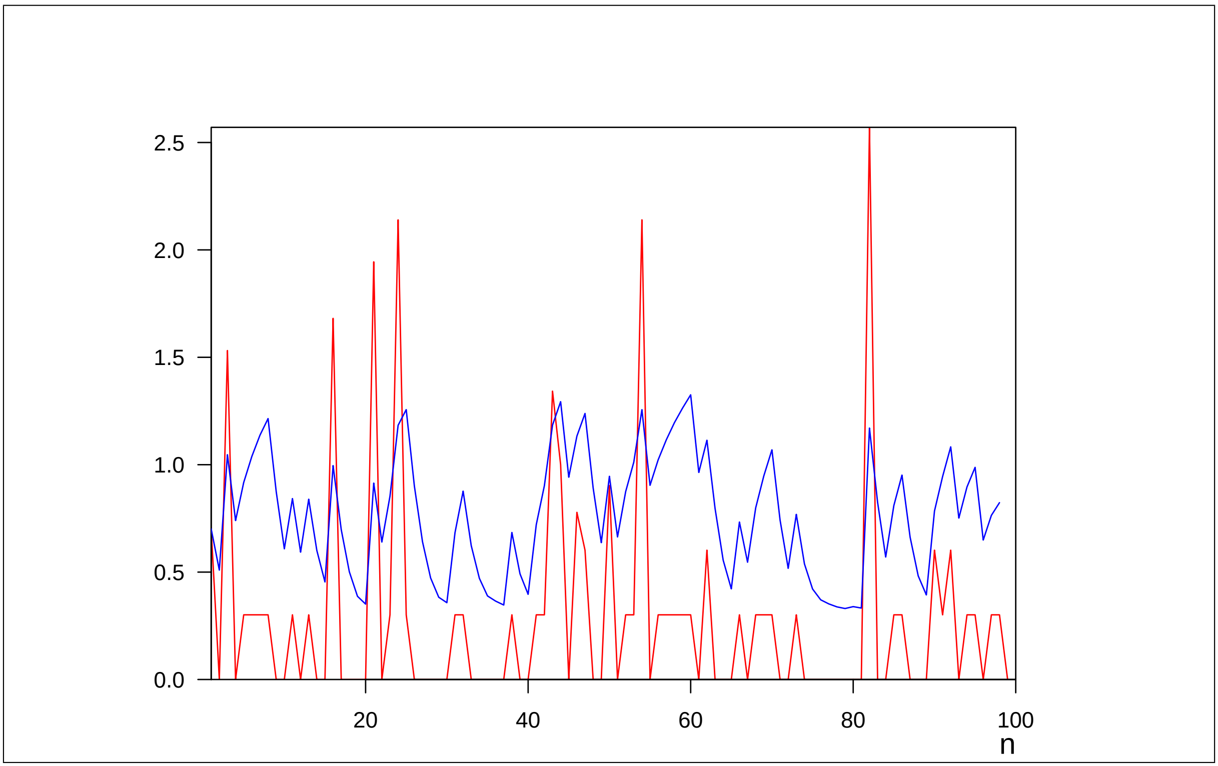


Figure 5.6: Example: $\log_{10}(Z)$, conditioned on $\{Z_{1000} > 0\}$, at the strictly descending ladder times of the associated random walk in Figure 5.5 (averaged over 1000 simulations) (blue) and $\log_{10} T^{(e)}$ of the excursion lengths $T^{(e)}$ between the strictly descending ladder points (red).

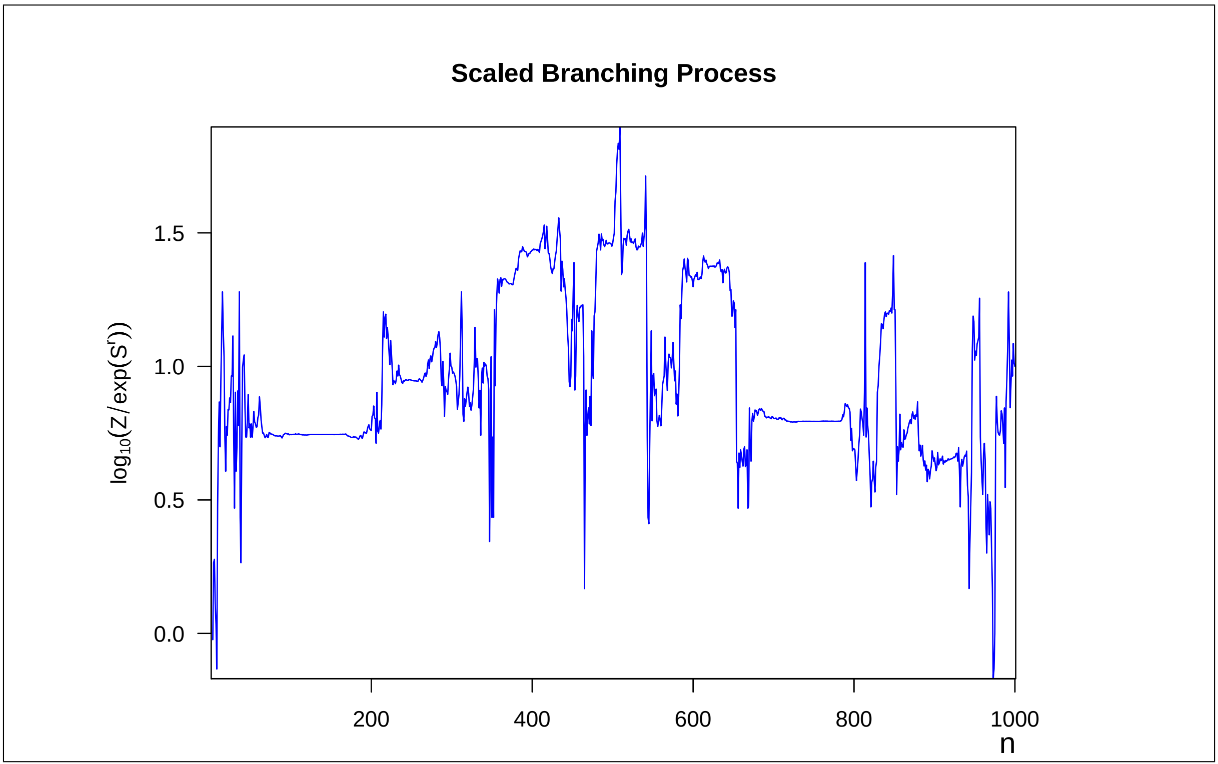


Figure 5.7: Example: $\log_{10}\left(Z/e^{S_r}\right)$, conditioned on $\{Z_{1000} > 0\}$ in the environment displayed in Figure 5.2.

Chapter 6

Perspectives

In this thesis, limit theorems describing intermediately subcritical BPREs have been developed. A more detailed describtion of the intermediately subcritical case is work in progress with Valery Afansayev, Götz Kersting and Vladimir Vatutin. For this purpose, the Geiger construction explained in Chapter 5 may be helpful. As already indicated by the simulations, the offspring numbers of the ancestral line from the Geiger construction is asymptotically size-biased distributed. Also, for any $\epsilon > 0$, the independent trees, conditioned on extinction before generation n that emerge to the left of this ancestral line before generation $\lfloor(1-\epsilon)n\rfloor$, are asymptotically unconditioned. This may allow a more detailed describtion of the bottlenecks (moments when the associated random walk is in a strict descending ladder point) as well as finding a proof for the conditional limit theorem claimed in the open problem in Chapter 3.3.

For general offspring distributions, the proper rate function for the upper large deviations of BPREs is described in detail in this thesis. The proof of the upper bound is however very technical. Finding a more natural and stochastic proof for the upper bound of the tail probabilities of Z_n, conditioned on the environment, is an open problem. Also Assumption $\mathcal{H}(\beta)$ may slightly be generalized, e.g. allowing the constants to depend on the environment.

Describing the lower large deviations of BPREs is work in progress with Vincent Bansaye. Here, the main task is finding the appropriate 'cost' of staying bounded, i.e. a representation of

$$\lim_{n\to\infty} -\frac{1}{n}\log \mathbb{P}(1 \le Z_n \le c) \quad , \; c > 0 \quad .$$

The quick simulation algorithm in Chapter 5 allows for a vizualization of conditioned BPREs, although only for geometric offspring distributions. This may also be helpful in applications like estimating the most recent common ancestor in a population.
A generalization to the linear fractional case, when the distribution of Z_n, conditioned on the environment, can be explicitely calculated, is possible if the focus is on the simulation of branching processes in varying environment. For more general offspring distributions, although there is an expression for the survival probability (see (3.13)), an explicit calculation of the conditioned distributions seems hardly feasible.

For the simulation of intermediately subcritical BPREs, conditioned on survival, there is the problem that – except for one-parametric offspring distributions determined by their expectation like the geometric distribution – the environment is no longer uniquely determined by the associated random walk. The convergence of the associated random walk then does not imply convergence of the environment any longer, conditioned on survival of the process. Finding the distribution of Π, conditioned on survival of Z_n, is an open problem.

Appendix A

Technical results

A.1 A general form of an inequality due to Paley and Zygmund

Here, a slight generalization of Lemma 4.1 in [Kal01] is proved:

Lemma A.1.1. *Let Y be a nonnegative random variable and assume $\mathbb{E}[|Y|^{1/p}] < \infty$ for some $p \in (0,1]$. Then for any $r \in (0,1)$ and $q = 1-p$,*

$$\mathbb{P}(Y \geq r\mathbb{E}[Y])^q \quad \geq \quad \frac{(1-r)\mathbb{E}[Y]}{\mathbb{E}[Y^{1/p}]^p} \ . \tag{A.1}$$

Proof. Assume $p, q \geq 0$ and $p+q = 1$. For any $c > 0$,

$$Y - c \quad \leq \quad Y\mathbf{1}_{\{Y>c\}} \ .$$

Taking the expectation and applying Hölder's inequality (see e.g. [Kal01, p. 15]),

$$\mathbb{E}[Y] - c \quad \leq \quad \mathbb{E}[Y^{1/p}]^p \mathbb{P}(Y > c)^q \ .$$

Setting $c = r\mathbb{E}[Y]$ yields the claim.

□

A.2 Slowly varying functions

In this section, we recall some properties of regularly varying functions. A standard reference is [BGT87]. In the following, let $\Upsilon : (0,\infty) \to (0,\infty)$ be a slowly varying function, that is for any $a > 0$,

$$\lim_{x\to\infty} \frac{\Upsilon(ax)}{\Upsilon(x)} = 1 \ .$$

A function $r : (0,\infty) \to (0,\infty)$ is called regularly varying with index α if for all $a > 0$

$$\lim_{x\to\infty} \frac{r(ax)}{r(x)} = a^\alpha \ .$$

Moreover this convergence holds uniformly with respect to a (see [BGT87, p. 22]) for $\alpha < 0$. For $\alpha = 0$, the convergence holds uniformly on each compact set.

A sequence $(b_n)_{n\geq 0}$ of positive numbers is called regularly varying with index α, if for any $a > 0$

$$\lim_{n\to\infty} \frac{b_{\lfloor an \rfloor}}{b_n} = a^\alpha \ .$$

Then the function $f(x) := b_{\lfloor x \rfloor}$ varies regularly with index α (see [BGT87, p. 52]).

Note that, by properties of slowly varying sequences (see [BGT87, Proposition 1.3.6, p. 16]), for any

$\delta > 0$, $z^{-\delta}\Upsilon(z) \to 0$ as $z \to \infty$. Therefore, for all $\delta > 0$ and z large enough, there is some constant C, such that $\Upsilon(z) \leq Cz^{\delta}$.

We use a Tauberian theorem for slowly varying functions from [Fel87, p. 423]. Here Γ denotes the Gamma-function (note that $\Gamma(\alpha+1)/\Gamma(\alpha+2) = 1/(\alpha+1)$). Also compare [BGT87, Theorem 1.5.11, p.28].

Theorem A.2.1. *Assume that $(b_k)_{k\geq 0}$ is a nonnegative, monotone sequence and $\alpha > -1$. Let $g(s) := \sum_{k=0}^{\infty} s^k b_k$ be convergent for $s \in [0,1)$ and Υ be slowly varying. Then*

$$b_k \sim \ k^{\alpha}\ \Upsilon(k) \quad (k \to \infty)\ ,$$

$$\sum_{k=1}^{n} b_k \sim \frac{1}{(\alpha+1)}\ n\ b_n\ ,$$

and

$$g(s) \sim \Gamma(\alpha+1)(1-s)^{-1-\alpha}\ \Upsilon\big(1/(1-s)\big) \quad (s \to 1-)$$

are equivalent.

Additionally, we need the following Lemma:

Lemma A.2.2. *Let $b_k := k^{\alpha}$. If $\alpha \geq -1$, then there is an $M < \infty$ such that for $s \in [0,1]$*

$$\sum_{k=1}^{\infty} s^k b_k \quad \leq \quad M\Upsilon(1/(1-s))(1-s)^{-1-\alpha}\ ,$$

where Υ is a slowly varying function (which is constant if $\alpha > -1$).

Proof. The function

$$\xi \quad := \quad s \to (1-s)^{1+\alpha} \sum_{k=1}^{\infty} s^k k^{\alpha}$$

is continuous on $[0,1)$.

For $\alpha > -1$, Theorem A.2.1 yields that

$$\lim_{s \to 1^-} \xi(s) \quad < \quad \infty$$

and thus ξ can be extended to a continuous function on $[0,1]$. Defining M as its supremum on $[0,1]$ proves the claim.

For $\alpha = -1$, $\sum_{k=1}^{\infty} \frac{s^k}{k} = -\log(1-s)$, which proves the lemma in this case as the logarithm is a slowly varying function. □

Remark. *For $\alpha < -1$,*

$$\sum_{j=k+1}^{\infty} j^{\alpha} \quad \leq \quad \int_k^{\infty} x^{\alpha} dx \ = \ \frac{k^{\alpha+1}}{-\alpha-1}\ . \tag{A.2}$$

A.3 Successive differentiation for the composition of functions

For the proof of the upper bound on the tail probabilities in Chapter 4.3 when $\beta > 2$, we need to calculate higher order derivatives of a composition of functions. Here, a useful formula for the l-th derivative of a composition of two functions, which could also be derived from the combinatorial form of Faà di Bruno's formula (see [dB55]), is proved.

Lemma A.3.1. *Let f and h be real-valued, l-times differentiable functions. Then*

$$\frac{d^l}{ds^l} h(f(s)) \quad = \quad \sum_{j=1}^{l} h^{(j)}(f(s)) u_{j,l}(s) \ , \tag{A.3}$$

where $u_{j,l}(s)$ is defined by

$$u_{j,l}(s) \quad := \quad \sum_{i=(i_1,\ldots,i_{2j})\in\mathcal{C}(j,l)} c_i (f^{(i_1)}(s))^{i_2} \cdots (f^{(i_{2j-1})})(s))^{i_{2j}} \tag{A.4}$$

with some constants $0 \leq c_i < \infty$ and $\mathcal{C}(j,l)$ defined by

$$\mathcal{C}(j,l) \quad := \quad \left\{ (i_1,\ldots,i_{2j}) \in \mathbb{N}^{2j} \middle| i_1 i_2 + i_3 i_4 + \ldots = l \ \textit{and} \ \ i_2 + i_4 + \ldots = j \right\} \ .$$

Proof. The formula is proved by induction with respect to l. For $l = 1$, by chain rule of differentiation, (A.3) is fulfilled. Assume that (A.3) and (A.4) hold for l. Then by product rule for differentiation,

$$\frac{d^{l+1}}{ds^{l+1}} h(f(s)) \quad = \quad \sum_{j=1}^{l} \left(h^{(j)}(f(s)) \frac{d}{ds} u_{j,l}(s) + u_{j,l}(s) f'(s) h^{(j+1)}(f(s)) \right) \ .$$

Thus

$$\begin{aligned} u_{j,l}(s) f'(s) \quad &= \quad \sum_{i\in\mathcal{C}(j,l)} c_i \big(f^{(1)}(s)\big)^1 (f^{(i_1)}(s))^{i_2} \cdots (f^{(i_{2j-1})}(s))^{i_{2j}} \\ &= \quad \sum_{i\in\mathcal{C}(j+1,l+1)} \tilde{c}_i (f^{(i_1)}(s))^{i_2} \cdots (f^{(i_{2j+1})}(s))^{i_{2(j+1)}} \ , \end{aligned}$$

with new constants defined by

$$\tilde{c}_{i_1,i_2,i_3,\ldots,i_{2(j+1)}} \quad := \quad \begin{cases} c_{i_3,\ldots,i_{2(j+1)}} & , \ \text{if } i_1 = i_2 = 1 \\ 0 & , \ \text{else} \end{cases} \ .$$

Furthermore,

$$\begin{aligned} \frac{d}{ds} u_{j,l}(s) \quad &= \quad \sum_{i\in\mathcal{C}(j,l)} \sum_{k=1}^{l} c_i (f^{(i_1)}(s))^{i_2} \cdots \ i_{2k} (f^{(i_{2k-1})}(s))^{i_{2k}-1} f^{(i_{2k-1}+1)}(s) \cdots (f^{(i_{2j-1})}(s))^{i_{2j}} \\ &= \quad \sum_{i\in\mathcal{C}(j,l+1)} \hat{c}_i (f^{(i_1)}(s))^{i_2} \cdots (f^{(i_{2j+1})}(s))^{i_{2(j+1)}} \ , \end{aligned}$$

again with some new constants $0 \leq \hat{c}_i < \infty$. This ends up the induction. □

List of Figures

Bibliography

[ABKV10] V.I. Afanasyev, C. Böinghoff, G. Kersting, and V.A. Vatutuin. Limit theorems for a weakly subcritical branching process in a random environment. *Journal of Theoretical Probability*, 2010.

[Afa80] V. I. Afanasyev. *Limit theorems for a conditional random walk and some applications.* Diss. cand. sci., MSU, Moscow, 1980.

[Afa93] V. I. Afanasyev. A limit theorem for a critical branching process in a random environment. *Diskret. Mat.*, 5(1):45–58, 1993.

[Afa97] V.I. Afanasyev. A new theorem for a critical branching process in random environment. *Discrete Math. Appl.*, 7:497–513, 1997.

[Afa98] V. I. Afanasev. Limit theorems for a moderately subcritical branching process in a random environment. *Diskret. Mat.*, 10(1):141–157, 1998.

[Afa01a] V. I. Afanasyev. On the maximum of a subcritical branching process in a random environment. *Stochastic Process. Appl.*, 93(1):87–107, 2001.

[Afa01b] V.I. Afanasyev. Limit theorems for an intermediately subcritical and a strongly subcritical branching processes in a random environment. *Discrete Math. Appl.*, 11(2):105–131, 2001.

[Afa10] V. I. Afanasyev. New invariance principles for critical branching process in random environment. In *Advances in data analysis*, Stat. Ind. Technol., pages 105–115. Birkhäuser Boston, Boston, MA, 2010.

[AGKV05a] V. I. Afanasyev, J. Geiger, G. Kersting, and V. A. Vatutin. Criticality for branching processes in random environment. *Ann. Probab.*, 33(2):645–673, 2005.

[AGKV05b] V. I. Afanasyev, J. Geiger, G. Kersting, and V. A. Vatutin. Functional limit theorems for strongly subcritical branching processes in random environment. *Stochastic Process. Appl.*, 115(10):1658–1676, 2005.

[Agr75] A. Agresti. On the extinction times of varying and random environment branching processes. *J. Appl. Prob.*, 12:39–46, 1975.

[AK71] K.B. Athreya and S. Karlin. On branching processes with random environments: I, II. *Ann. Math. Stat.*, 42:1499–1520, 1843–1858, 1971.

[AN72] K.B. Athreya and P. E. Ney. *Branching Processes.* Dover Publications, New York, 1. edition, 1972.

[Ban09] V. Bansaye. Surviving particles for subcritical branching processes in random environment. *Stochastic Process. Appl.*, 119(8):2436–2464, 2009.

[BB93] J. D. Biggins and N. H. Bingham. Large deviations in the supercritical branching process. *Adv. in Appl. Probab.*, 25:757–772, 1993.

[BB09] V. Bansaye and J. Berestycki. Large deviations for branching processes in random environment. *Markov Process. Related Fields*, 15(4):493–524, 2009.

[BB10] V. Bansaye and C. Böinghoff. Upper large deviations for branching processes in random environment with heavy tails. preprint: arXiv:1004.1263v1, 2010.

[BD94] J. Bertoin and R. A. Doney. On conditioning a random walk to stay nonnegative. *Ann. Probab.*, 22(4):2152–2167, 1994.

[BDKV10] C. Böinghoff, E. Dyakonova, G. Kersting, and V.A. Vatutin. Branching processes in a random environment which extinct at a given moment. *Markov Process. Relat. Fields*, 16:329–350, 2010.

[BGK05] Matthias Birkner, Jochen Geiger, and Götz Kersting. Branching processes in random environment—a view on critical and subcritical cases. In *Interacting stochastic systems*, pages 269–291. Springer, Berlin, 2005.

[BGT87] Bingham, Goldie, and Teugels. *Regular Variation.* Cambridge University Press, Cambridge, 1. edition, 1987.

[BK10] C. Böinghoff and G. Kersting. Upper large deviations for branching processes in a random environment - offspring distributions with geometrically bounded tails. *Stoch. Proc. Appl.*, 120:2064–2077, 2010.

[BV97] K. A. Borovkov and V. A. Vatutin. Reduced critical branching processes in random environment. *Stochastic Process. Appl.*, 71(2):225–240, 1997.

[CT84] J. Coffey and D. Tanny. A necessary and sufficient condition for noncertain extinction of a branching process in a random environment (BPRE). *Stochastic Process. Appl.*, 16(2):189–197, 1984.

[dB55] F. Faá di Bruno. Sullo sviluppo delle funzioni. *Annali di Scienze Matematiche e Fisiche*, 6, 1855.

[DGV04] E. E. Dyakonova, J. Geiger, and V. A. Vatutin. On the survival probability and a functional limit theorem for branching processes in random environment. *Markov Process. Related Fields*, 10(2):289–306, 2004.

[DH97a] J. C. D'Souza and B. M. Hambly. On the survival probability of a branching process in a random environment. *Adv. in Appl. Probab.*, 29(1):38–55, 1997.

[DH97b] J. S. D'Souza and B.M. Hambly. On the survival probability of a branching process in a random environment. *Adv. Appl. Probab.*, 29, 1997.

[dH00] Frank den Hollander. *Large Deviation.* American Mathematical Society, Providence, RI, 1. edition, 2000.

[DIM77] R. T. Durrett, D. L. Iglehart, and D. R. Miller. Weak convergence to Brownian meander and Brownian excursion. *Ann. Probability*, 5(1):117–129, 1977.

[Don77] R. A. Doney. A note on a condition satisfied by certain random walks. *J. Appl. Probability*, 14(4):843–849, 1977.

[Don85] R. A. Doney. Conditional limit theorems for asymptotically stable random walks. *Z. Wahrsch. Verw. Gebiete*, 70(3):351–360, 1985.

[Dur78] R. Durrett. Conditioned limit theorems for some null recurrent Markov processes. *Ann. Probab.*, 6(5):798–828, 1978.

[DZ93] A. Dembo and O. Zeitoni. *Large Deviations Techniques and Applications.* Jones and Barlett Publishers International, London, 1. edition, 1993.

[Fel68] William Feller. *An Introduction to Probability Theory and Its Applications- Volume I.* John Wiley & Sons, Inc., New York, 3. edition, 1968.

[Fel87] William Feller. *An Introduction to Probability Theory and Its Applications.* John Wiley & Sons, Inc., New York, 2. edition, 1987.

[Gei99] Jochen Geiger. Elementary new proofs of classical limit theorems for Galton-Watson processes. *J. Appl. Probab.*, 36(2):301–309, 1999.

[GK00] J. Geiger and G. Kersting. The survival probability of a critical branching process in random environment. *Teor. Veroyatnost. i Primenen.*, 45(3):607–615, 2000.

[GKV03] Jochen Geiger, Götz Kersting, and Vladimir A. Vatutin. Limit theorems for subcritical branching processes in random environment. *Ann. Inst. H. Poincaré Probab. Statist.*, 39(4):593–620, 2003.

[GL01] Yves Guivarc'h and Quansheng Liu. Propriétés asymptotiques des processus de branchement en environnement aléatoire. *C. R. Acad. Sci. Paris Sér. I Math.*, 332(4):339–344, 2001.

[Gui85] Y. Guivarc'h. *Processus de branchement en environnement aléatoire*. Publications de l'Institut de Recherche Mathématique de Rennes. [Publications of the Rennes Institute of Mathematical Research]. Université de Rennes I Institut de Recherche Mathématique de Rennes, Rennes, 1985.

[GZ91] D. R. Grey and Lu Zhunwei. Extinction probabilities of branching processes in random environments. In *Selected Proceedings of the Sheffield Symposium on Applied Probability (Sheffield, 1989)*, volume 18 of *IMS Lecture Notes Monogr. Ser.*, pages 205–211. Inst. Math. Statist., Hayward, CA, 1991.

[Ham92] B. Hambly. On the limiting distribution of a supercritical branching process in a random environment. *J. Appl. Probab.*, 29(3):499–518, 1992.

[HL10] C. Huang and Q. Liu. Moments, moderate and large deviations for a branching process in a random environment. preprint: arXiv:1007.1738v1, 2010.

[Kal01] Olav Kallenberg. *Foundations of Modern Probability.* Springer, London, 2. edition, 2001.

[Koz76] M. V. Kozlov. The asymptotic behavior of the probability of non-extinction of critical branching processes in a random environment. *Teor. Verojatnost. i Primenen.*, 21(4):813–825, 1976.

[Koz95] M. V. Kozlov. A conditional functional limit theorem for a critical branching process in a random environment. *Dokl. Akad. Nauk*, 344(1):12–15, 1995.

[Koz06] M. V. Kozlov. On large deviations of branching processes in a random environment: geometric distribution of descendants. *Discrete Math. Appl.*, 16(2):155–174, 2006.

[Liu96] Q. Liu. On the survival probability of a branching process in a random environment. *Ann. Inst. H. Poincaré Probab. Statist.*, 32(1):1–10, 1996.

[Nav97] Philippe Naveau. Comparison between the Chernoff and factorial moment bounds for discrete random variables. *Amer. Statist.*, 51(1):40–41, 1997.

[Pro00] Proceedings of the 9th International Summer School on Probability Theory and Mathematical Statistics. *Large deviations and branching processes*, volume 13. Pliska Stud. math. Bulgar., 2000.

[RW00] L.C.G. Rogers and David Williams. *Diffusions, Markov Processes, and Martingales-Ito Calculus*, volume 2. Cambridge University Press, Cambridge, 2. edition, 2000.

[Smi68] Walter L. Smith. Necessary conditions for almost sure extinction of a branching process with random environment. *Ann. Math. Statist*, 39:2136–2140, 1968.

[Spi64] F. Spitzer. *Principles of Random walks.* Van Nostrand, Princeton, 1. edition, 1964.

[SW69] Walter L. Smith and William E. Wilkinson. On branching processes in random environments. *Ann. Math. Statist.*, 40:814–827, 1969.

[Tan77] D. Tanny. Limit theorems for branching processes in a random environment. *Ann. Probability*, 5(1):100–116, 1977.

[Tan88] D. Tanny. A necessary and sufficient condition for a branching process in a random environment to grow like the product of its means. *Stochastic Process. Appl.*, 28(1):123–139, 1988.

[Tan78] D. Tanny. Normalizing constants for branching processes in random environments (B.P.R.E.). *Stochastic Process. Appl.*, 6(2):201–211, 1977/78.

[Vat04] V.A. Vatutin. Limit theorem for an intermediate subcritical branching process in a random environment. *Theory Probab. Appl.*, 48(3):481–492, 2004.

[VD97] V. A. Vatutin and E. E. Dyakonova. Critical branching processes in a random environment: probabilities of extinction at a fixed time. *Diskret. Mat.*, 9(4):100–126, 1997.

[VD02] V. A. Vatutin and E. Dyakonova. Reduced branching processes in random environment. In *Mathematics and computer science, II (Versailles, 2002)*, Trends Math., pages 455–467. Birkhäuser, Basel, 2002.

[VD04] V.A. Vatutin and E. Dyakonova. Yaglom type limit theorem for branching processes in random environment. In *Mathematics and computer science. III*, Trends Math., pages 375–385. Birkhäuser, Basel, 2004.

[VK08] V. A. Vatutin and A. E. Kyprianou. Branching processes in random environment die slowly. In *Fifth Colloquium on Mathematics and Computer Science*, Discrete Math. Theor. Comput. Sci. Proc., AI, pages 375–395. Assoc. Discrete Math. Theor. Comput. Sci., Nancy, 2008.